15 0116.1

→ See p. 16. 2d ed.

STUDIES

FROM THE

Yale Psychological Laboratory

EDITED BY

EDWARD W. SCRIPTURE, PH.D.

Instructor in Experimental Psychology

Volume II.

Issued on November 1st, 1894

CONTENTS

YALE UNIVERSITY, NEW HAVEN, CONN.

Price, $1. Obtainable at NEW HAVEN, Psychological Laboratory, 109 Elm Street; NEW YORK, Charles Scribner's Sons, 153 Fifth Ave.; Gustav E. Stechert, 810 Broadway; BERLIN, Mayer & Müller, Markgrafen-Strasse 51; LEIPZIG, Bernhard Liebisch, Kurprinz-Strasse 6; PARIS, E. Leroux, 28, Rue Bonaparte; BOLOGNA, Pietro Virano, Piazza Galvani.

STUDIES

FROM THE

Yale Psychological Laboratory

EDITED BY

EDWARD W. SCRIPTURE, PH.D.
Instructor in Experimental Psychology

VOL. II
Issued on November 1st, 1894

YALE UNIVERSITY,
NEW HAVEN, CONN.
1894

THE TUTTLE, MOREHOUSE & TAYLOR PRESS

CONTENTS.

ON MEAN VALUES FOR DIRECT MEASUREMENTS,

BY

E. W. Scripture.

Sources of error.

In making measurements we are subject to errors which can be classed as: 1. errors of scale, 2. errors of observation, 3. errors of definition, 4. errors of number, 5. errors of calculation.

1. *Errors of scale.* Since the divisions of the standard used in measuring come into consideration, it must be determined on every occasion what aliquot portion of the unit, or what ultimate sub-unit shall or can be finally considered. Let this sub-unit be r. On the axis of X (fig. 1) let the value of the quantity measured be

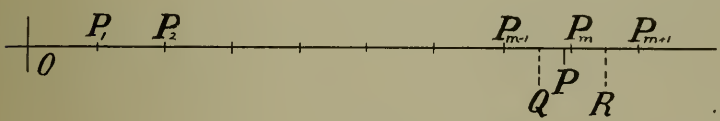

Fig. 1.

at P when expressed with infinite accuracy. Let the sub-unit be applied in succession from O, producing a series of points P_1, P_2, . . ., P_{m-1}, P_m, P_{m+1}. If P_m is the nearest point to P, the number mr is noted as the value of P. The measurement thus consists not in giving the absolutely true value of P but in noting the nearest mark of the measuring instrument. The error of scale $P-P_m$ in any particular case cannot be known, as all the points between Q and R will be denoted by the same result. It is usually assumed that results between $Q = P_m + \frac{r}{2}$ and $R = P_m - \frac{r}{2}$ are equally probable and that in the long run the result $P_m = mr$ will differ but little from P.

The assumption of equal probability for an error of $+\frac{r}{2}$ and $-\frac{r}{2}$ is not strictly correct. Particularly wrong is the further assumption that all values between $+\frac{r}{2}$ and $-\frac{r}{2}$ are equally probable.

If all other sources of error be made negligible, i. e. if the errors of observation and of definition be made sufficiently small, the re-

sulting value P can be determined in each case with an approximation so close that it can be regarded, comparatively, as the true value. The scale-readings for a set of results can then be compared with the "true" values. Thus a series of differences

$$P^{(1)}-P_m=V_1;\ P^{(2)}-P_m=V_2;\ .\ \ .\ \ .;\ P^{(n)}-P_m=V_n$$

are obtained, which can be called the errors of scale.

The real values of the measurements are thus distributed over a region $P^{(1)}-P^{(n)}$ for which one value P_m is read as a representative. It thus becomes a question how well P_m represents the true values.

The influences that affect the error of scale, whereby the actual reading may be too great or too small, may be said to be of two kinds: 1. instrumental, 2. psychological.

An example of the former may be seen in all instruments in which the reading depends on a bar dropped into the teeth of a rack, as in the Hipp chronoscope. If the edges of the bar and of the teeth were infinitely sharp, the bar would (other sources of error supposed absent) drop and slide down just as often on one side of a tooth as the other. If, however, the edges are rounded, as they must eventually become by wear, the inertia of the body in motion, whether the bar or the rack, will carry the bar constantly to one side whenever it strikes on the rounded edge. Thus in the Hipp chronoscope the readings are all slightly too large. A reading of m^σ is supposed to represent all values between $(m+\frac{1}{2})^\sigma$ and $(m-\frac{1}{2})^\sigma$ whereas in an old instrument m^σ may be the reading possibly for $(m+\frac{1}{4})^\sigma$ to $(m-\frac{3}{4})^\sigma$.

An example of the latter is found in the familiar case of reading in tenths of the index-unit on a graduated scale, as in getting thousandths of a second from a tuning-fork curve in hundredths. A large portion of the work on the least perceptible difference might be used to determine the law of frequency for the error of scale.

The assumption of equal probability for an error of $+\frac{r}{2}$ and $-\frac{r}{2}$ is not strictly correct, as the law of probability followed by the measurements in general will be followed here also. It has been proven, however, that the error introduced by the assumption is of the second order of smallness as compared with the error of scale.[1]

2. *Errors of observation.* In making measurements the error of

---

[1] Lehmann-Filhés, *Ueber Ausgleichung abgerundeter Beobachtungen*, Astr. Nachr., 1889 CXX 305.

Pizzetti, *Sur la théorie des observations arrondies*, Astr. Nachr, 1890 CXXIV 33.

scale can be made so small as to be negligible in comparison with all other sources of error. The measurements $x_1, x_2, \ldots, x_n$ can be considered as having been recorded with practically perfect accuracy. They will disagree owing to sources of error thus classified by LAMBERT:[1] inaccuracy of the instruments, lack of care of the observer, dullness of the senses, and circumstances of the observation. The errors of observation resulting herefrom falsify the separate measurements. As these sources of error are made smaller the values of x differ less from one another and we can suppose that they would ultimately tend toward a value X which is usually called the true value of x. If we could know this true value, the separate errors of observation would be given by

$$V_1=x_1-X;\ V_2=x_2-X;\ \ldots;\ V_n=x_n-X.$$

As we cannot know X we cannot know the values of V. If we take some representative value R, we obtain

$$v_1=x_1-R;\ v_2=x_2-R;\ \ldots;\ v_n=x_n-R.$$

The best we can do is to determine that the value R shall not differ from X by more than a given amount. This is the problem of adjustment of measurements as it is usually proposed in works on physical measurements.

3. *Errors of definition.* When the errors of scale and of observation are made so small as to be entirely negligible, the results $x_1, x_2, \ldots, x_n$ will be practically true as far as these errors are concerned. They will still not agree owing to the infinite number of factors entering into the definition of the quantity. No matter how carefully we define it, we can never make it absolutely complete.

Let the quantity measured be the height of the American soldier. The object is already limited to a nation, a sex, a class and a range of ages. Let the unit of scale be 0.001^{m} and the method of measurement be a blunt point descending on an arm supported without shake by a rigid vertical bar. The recorded heights will extend over a range of 0.60^{m}. The height of the American soldier is thus uncertain to that amount.

Let the quantity be further limited to a given age, say 30 yrs., and a given nativity, say Massachusetts ; this may render the result definite to 0.30^{m}. Let it be further confined to a given company of a given Massachusetts regiment at a given day. The result might well be uncertain to 0.15^{m}.

[1] LAMBERT, *Theorie d. Zuverlässigkeit d. Beobachtungen u. Versuche*, Beyträge z. Gebr. d. Math., I 424.

These limitations are still indefinite. As the age may range over 12 months and the constituency of the company might change, let the measurement be limited to a given individual, and let 10 measurements be made in the manner prescribed. The results will depend on the altitude and inclination of the head. Even if these be defined as the extreme height while standing on the heels, the individual will, owing to practice and to fatigue, never stand twice alike. Let the required height be the maximum attainable under any conditions; the results will vary unless the occasion be fixed at a given time. Although the subject has never been investigated, the varying atmospheric conditions during the day undoubtedly affect the activity of the nervous system and thus the tension of the muscles maintaining the upright position. Even when limited to a definite occasion the results will vary unless the pressure of the point on the skin be defined. Here, however, the limits of accuracy of the proposed method of observation will be reached; there will probably be no changes comparable to the inaccuracy of the observer's eye in adjusting the point. A finer method with multiplying levers and air-transmission by Marey capsules will reveal continual, though minute, fluctuations. The error of definition could in this manner be reduced probably to 0.0001^{m} by taking the height thus indicated at a given instant. With still finer methods familiar to physicists the instant of time would necessarily be more closely defined. Enough, however, has been said to show that owing to the impossibility of defining with infinite accuracy the quantity measured, the true value of the quantity is to be considered as that which is obtained with an accuracy sufficient for the purpose in hand; whether it be a general figure by which to compare the American with the French soldier, by which to compare those from the different states, by which to designate a particular individual, or by which to determine the individual's change at each instant.

Let it be required to determine the reaction-time of a given individual on a given occasion to a given stimulus of a given intensity. Let the external conditions be further defined by perfect quietness and perfect darkness; let the air-supply be of a given quality, the pulse-rate of a given frequency, etc. Let the method of recording be accurate to 1^{σ}, i. e. the error of observation shall be less than that amount. Even under these circumstances the results will vary owing to the continually changing subjective conditions of attention, fatigue, etc., which are still beyond control. It is only in exceptional cases that the disagreement can be reduced below an average

of 10^σ. As the methods of psychology are perfected we shall be able to define and control each of the subjective conditions influencing the time, until we can define the reaction-time so that it shall not vary to the extent of 1^σ under the given conditions.[1] Of course, when this occurs the errors of recording must be proportionately small.

When the error of definition is negligible in comparison with the errors of observation in direct measurements, physicists speak of the "true" value of the quantity which would be attainable with an infinitely accurate method of observation. As physicists are able in many fundamental cases, e. g. length, time, etc., to define more accurately than they can observe, the "true" value in this sense is often sharply distinguished from an average value. There is a "true" value for the height of the barometer for any given minute because we cannot observe finely enough to detect the continuous fluctuation.

Enough has been said, I think, to indicate that the distinction sometimes made between the average of observations and the mean of a series of quantities[2] is really a distinction between the mean of a series of quantities subject to variation in observation and the mean of a series of quantities subject to variation in definition.[3]

4. *Errors of number.* The result of a measurement is recorded with a definite number of decimal places, whereas a perfectly true value would require an infinite number. The possible numerical error of any result recorded to α places is 0.5β, where β is 1 unit in the α place. Owing to the partial cancellation of these errors in combining measurements, the mean numerical error in the α place is 0.25β.

5. *Errors of calculation.* The errors caused by mistakes in reading scale-numbers, in writing them down and in performing the necessary arithmetical calculations can be treated by the principles of probability.[4] For the sake of simplicity they will be left out of consideration here.

---

[1] SCRIPTURE, *Accurate work in psychology*, Am. Jour. Psych., 1894 VI 427.

[2] QUETELET, Théorie des prob., Lettre XI, Bruxelles 1846.

HERSCHEL, QUETELET *on probabilites*, Edinburgh Rev., 1850 CLXXXV 1; Essays, 365 (404), Lond 1857; also in QUETELET, Physique sociale de l'homme, I 35, Bruxelles 1869.

JEVONS, Principles of Science, 2. ed., 362, Lond. 1887.

[3] VENN, Logic of Chance. 2. ed., 96, Lond. 1876.

[4] HOFMANN, *Ermittelung der Tragweite der Neunerprobe bei Kenntniss der subj. Genauigkeit des Rechnenden*, Zt. f. Math. u. Phys., 1889 XXXIV 116.

EMMERICH, *Zur Neunerprobe*, Zt. f. Math. u. Phys., 1889 XXXIV 320.

Representative values.

Owing to the various sources of error the measurements x_1, x_2, . . ., x_n will generally disagree. The disagreeing results of the measurements, regarded in themselves, are actual concrete matters of fact. For practical reasons these values are replaced by one value deduced from them on some given principle. The proper selection of the representative value requires a clear conception of the purpose for which it is to be used and of the method of deduction.

It is sometimes desirable to choose an extreme value to represent the actual ones, e. g. the highest mast to be allowed for in building a bridge, the smallest difference that might be noticed, the largest variation for a given degree of probability; but for the purpose here under consideration some mean value is desired and to this the discussion will be confined.

Proposed means.

Means have been proposed on four principles: 1. on consideration of the properties of a geometrical or material figure employed in representing the results and their probabilities; 2. on the principles of probability *a priori;* 3. on the possible ways of practically calculating means; 4. on certain relations of the powers of the variations.

1. *General analytical or geometrical deduction.* If we denote the number of occurrences of the result x_k by m_k, where $k = 1, 2, \ldots, n$, then the quotients

$$\frac{m_1}{m}, \quad \frac{m_2}{m}, \quad \ldots, \quad \frac{m_n}{m},$$

where $m = m_1 + m_2 + \ldots + m_n$, will denote the relative frequencies of x_1, x_2, . . ., x_n respectively. These frequencies stand in the relation

$$m_1 : m_2 : \ldots : m_n.$$

When the individuals are taken at random, as m is made greater, the quotients tend toward definite limits,

$$\phi(x_1), \phi(x_2), \ldots, \phi(x_n),$$

known as the probabilities of x_1, x_2, . . ., x_n respectively.

a. Selection on the basis of $y = \phi(x)$. Let the given values be laid off on the axis of X (fig. 2) and the ordinates y_1, y_2, . . ., y_n,

be erected proportionately to $\phi(x_1)$, $\phi(x_2)$, . . ., $\phi(x_n)$. For continuous values,

$$y = \phi(x) \tag{1}$$

will be the expression for the curve or law of frequency or probability.

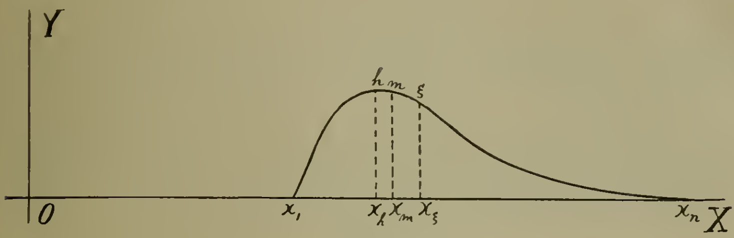

FIG. 2.

The probability of any value x may be regarded as the probability of a value falling between x and $x + dx$. This will be represented by the area inclosed by the curve over the base dx with the mean ordinate $\phi(x)$. The total area is

$$W = \int_{x_1}^{x_n} \phi(x) dx. \tag{2}$$

Since the probability of a result outside of the given limits is not 0 but is infinitely small, it is justifiable to write

$$W = \int_{-\infty}^{+\infty} \phi(x) dx. \tag{3}$$

The form (2) will, however, be retained here as the extension of the limits has occasioned some misunderstanding.[1]

Since it is certain that all values are included between $-\infty$ and $+\infty$,

$$W = 1. \tag{4}$$

The value of maximum probability is the abscissa x_h of the highest point of the curve. It is found by putting

$$\frac{d\phi(x)}{dx} = 0$$

and taking that one of the resulting values for which $d^2\phi(x)/dx^2$ is negative.

---

[1] CATTELL, *On errors of observation*, Am. Jour. Psych., 1893 V 287.

The value of mean area x_m is the abscissa whose ordinate divides the area of the curve into two equal parts, and is found from

$$\int_{x_1}^{x_m} \phi(x)dx = \int_{x_m}^{x_n} \phi(x)dx = \tfrac{1}{2}W = \tfrac{1}{2}. \qquad (5)$$

This value of mean area was called by LAPLACE[1] the value of middle probability. For the sake of brevity the name can be shortened into "middle value."

The probabilities $\phi(x_1), \phi(x_2), \ldots, \phi(x_n)$ can be regarded as parallel forces acting on the points $x_1, x_2, \ldots, x_n$ of a straight line. The centroid or mean center will be at

$$x_\xi = \frac{\Sigma\phi(x)\,x}{\Sigma\phi(x)} \qquad (6)$$

If $\phi(x_1), \phi(x_2), \ldots, \phi(x_n)$ be regarded as masses at the points of a straight line, the position of the center of gravity will be expressed by the same equation.

For continuous values,

$$x_\xi = \frac{\int_{x_1}^{x_n} x\phi(x)dx}{\int_{x_1}^{x_n} \phi(x)dx}; \qquad (7)$$

or, on account of (2) and (4),

$$x_\xi = \int_{x_1}^{x_n} x\phi(x)dx. \qquad (8)$$

This equation also represents the abscissa of the center of gravity of an area of uniform density bounded by the axis of X and the curve $y = \phi(x)$.

These values are represented in fig. 2. The highest point of the curve is at h; its ordinate is x_h. The ordinate m, x_m divides the area of the curve into halves. The mean center or center of gravity is on ξ; its abscissa is x_ξ.

b. Selection on the basis of $Y = \int\phi(x)dx$. Starting with x_1 on the axis of X, erect the ordinate $Y_1 = \phi(x_1)$. At $x = x_2$, erect $Y_2 = \phi(x_1) +$

---

[1] LAPLACE, *Mémoire sur la probabilité des causes par les évènements*, Mém. de math. et de phys. par divers savans, Acad. Paris, 1774 VI 621 (636).

$\phi(x_2)$; at x_k, ($k=1, 2, \ldots, n$) erect $Y_k=\phi(x_1)+\phi(x_2)+\ldots+\phi(x_k)$. The unit of abscissa is here, as before, dx, and just as in the previous case, these values can be transformed into continuous ones. Thus

$$Y_k=\int_{x_1}^{x_k}\phi(x)\,dx. \tag{9}$$

This is the integral curve for $y=\phi(x)$, and may be plotted or drawn directly from it.[1]

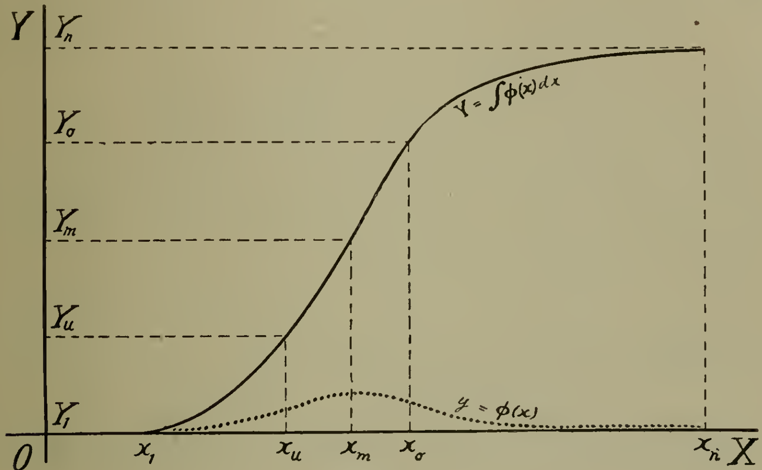

FIG. 3.

The difference between the ordinates at the beginning and end of the curve will correspond to the total area of the frequency curve. Thus

$$Y_n-Y_1=\int_{x_1}^{x_n}\phi(x)\,dx=W=1.$$

The value x_m whose ordinate Y_m is halfway between Y_1 and Y_n is determined by

$$Y_n-Y_m=Y_m-Y_1$$

or

$$\int_{x_1}^{x_m}\phi(x)\,dx=\int_{x_m}^{x_n}\phi(x)\,dx=\tfrac{1}{2}W=\tfrac{1}{2}. \tag{10}$$

This value x_m is evidently the value of middle probability noted above.

[1] ABDANK-ABAKANOWICZ, *Les integraphes, la courbe intégrale et ses applications*, Paris 1886.

The over-quartile x_o and the under-quartile x_u are determined by

$$x_o = f\left(\frac{3(Y_n - Y_1)}{4}\right), \quad x_u = f\left(\frac{Y_n - Y_1}{4}\right)$$

The octiles are determined by

$$x_k = f\left(\frac{k(Y_n - Y_1)}{8}\right), \quad (k = 1, 2, \ldots, 8),$$

These may be used as characteristic values to indicate the form of the curve.[1]

The percentiles

$$x_c = f\left(\frac{c(Y_n - Y_1)}{100}\right), \quad (c = 1, 2, \ldots, 100)$$

have formed the basis of the method of percentile grades.[2] In practice the percentiles generally become vigintiles or deciles.

Galton, who first introduced the practical use of this method of considering measurements, treats them in a way which, mathematically stated, is as follows. With any arbitrary unit on the axis of X, erect in succession the ordinates

$$Y_1 = x_1; \; Y_2 = x_1; \; \ldots; \; Y_{m'} = x_1,$$

extending the processes till the value x_1 has been used as many times as it occurs in the set of results. Likewise let

$$Y_{m'+1} = x_2; \; Y_{m'+2} = x_2; \; \ldots, \; Y_{m'+m''} = x_2,$$

where x_2 has occurred m'' times. Repeating this process we have

$$Y_{m'+m''+1} = x_3; \; Y_{m'+m''+2} = x_3; \; \ldots; \; Y_{m'+m''m'''} = x_3;$$

. . . .

$$Y_{m'+m''+ \ldots +m_{r-1}+1} = x_r; \; Y_{m'+m''+ \ldots +m_{r-1}+2} = x_r; \; \ldots;$$
$$Y_{m'+m''+ \ldots +m_{r-1}+m_r} = x_r;$$

. . . .

[1] Galton, Hereditary Genius, 33, London 1870.

Galton, *Statistics by intercomparison*, Phil. Mag., 1875 (4) XLIX 33.

Galton, Inquiries into Human Faculty, 51, New York 1883.

McAlister, *The law of the geometric mean*, Proc. Royal Soc. London, 1879 XXIX 367 (374).

Edgeworth, *Problems in probabilities*, Phil. Mag., 1886 (5) XXII 371 (374).

[2] Galton, Natural Inheritance, London 1889. (I have not seen this work.)

Bowditch, *Physique of women in Massachusetts*, XXI. Ann. Rept. Mass. State Board of Health, 285, Boston 1890.

Bowditch, *Growth of children*, XXII. Ann. Rept. Mass. State Board of Health, 479, Boston 1891.

Seaver, Manual of Anthropometry, 1. chart, New Haven 1890.

Geissler, *Ueber die Vorteile der Berechnung nach perzentilen Graden*, Allg. stat. Arch., 1891–1892 452.

$$Y_{m'+m''+\cdots+m_{n-1}+m}=x_n;\quad Y_{m'+m''+\cdots+m+_{n-1}+3}=x_n;\quad \cdots;$$
$$Y_{m'+m''+\cdots+m_{n-1}+m_n}=x_n.$$

Thus in general Y_{m_k} is the ordinate for

$$X_{m_k}=\sum_{m'}^{m_k} m,\quad (k=1, 2, \ldots, n).$$

By considering the whole interval covered by X to be $m=m'+m''+\ldots+m_n$, with the sub-intervals $\frac{m'}{m}, \frac{m''}{m}, \ldots, \frac{m_n}{m}$ we have for any point X_k

$$X_k=\sum_{m'}^{m_k}\frac{m_i}{m},\quad (i=1, 2, \ldots, k).$$

For continuous values this evidently becomes of the same form as (9) and would become exactly the same if the Y-axis had been used in place of the X-axis. GALTON's elementary illustration[1] of a group of men placed side by side in order of size with a curve just touching the tops of their heads has led to the use of the axis of Y for the values of $x_1, x_2, \ldots, x_n$. As GALTON's method and illustration have been widely accepted by statisticians, confusion is introduced by the neglect of mathematical conventions. To the non-mathematical statistician it seems more natural to erect height-ordinates vertically and to imagine a row of men standing on their feet. He should remember, however, that in the simple probability curve the heights were laid off horizontally on the axis of X and that if he wishes to have the height-ordinates vertical they must in both cases be laid off on the axis of Y. GALTON's "ogive" curve is obtained by tracing the integral curve on tissue-paper, looking at it from the back of the paper and turning it through 90°. But if this is done, the simple probability-curve must be treated in the same way. In respect to the proper choice of axes SEAVER's table, for example, is correct, BOWDITCH's tables and curves are not.

2. *The most probable value.* If x be taken to represent a set of values $x_1, x_2, \ldots, x_n$, the differences x_k-x_p, $(k=1, 2, \ldots, n)$, can be considered as errors or detriments.[2]

The antecedent probability of any set of errors is proportional to

$$\phi(x_1-x_p)\ \phi(x_2-x_p)\ \ldots\ \phi(x_n-x_p).$$

[1] GALTON, *Statistics by intercomparison*, Phil. Mag., 1875 (4) XLIX 33.

[2] GAUSS, Theoria combinationis observationum, I, 6.

That value of x_p which renders this product a maximum is *a priori* the most probable value. This requires that

$$0 = \frac{1}{\phi(x_1 - x_p)} \frac{d\phi(x_1 - x_p)}{d(x_p)} + \frac{1}{\phi(x_2 - x_p)} \frac{d\phi(x_2 - x_p)}{d(x_p)} + \\ \quad . \quad . \quad . + \frac{1}{\phi(x_n - x_p)} \frac{d\phi(x_n - x_p)}{d(x_p)},$$

which is the equation of condition for the determination of x_p.[1]

This equation is used in two ways: 1. to determine the law of probability required for an assumed most probable value; 2. to determine the most probable value for an assumed law of probability. In either case an arbitrary assumption must be made, as GAUSS clearly recognized.[2] If the average be assumed as the most probable value, then, with the usual additional assumptions,

$$\phi(v) = \frac{h}{\sqrt{\pi}} e^{-h^2 v^2} \tag{11}$$

where $v_k = x_k - A, \quad (k = 1, 2, \ . \ \ . \ \ ., n)$

and $A = \frac{\Sigma x}{n}.$

If another mean, e. g. the median or the geometric mean, be assumed as the most probable value, the law of frequency takes a different form.

It has been customary to regard the assumption of the arithmetical mean and the exponential law of error therefrom deduced, as practically verified, although theoretically not correct.[3]

Since the fundamental supposition of symmetry of the probability-curve is quite unjustifiable for psychological measurements, and since the theory of means has been treated by GAUSS independently of any definite law of error, it is justifiable to omit all further consideration of this treatment of the most probable value, although GAUSS's earlier treatment is followed by most of the text-books on the adjustment of measurements.[4]

---

[1] GAUSS, Theoria motus corp. coel., II, 3, 177.

[2] GAUSS, Theoria motus corp. coel., II, 3, 177.
GAUSS, *Anzeige*, Gött. gel. Anz., 1821 Feb. 26; Werke, IV 98.

[3] BERTRAND, Calcul des probabilités, 183, Paris 1889.

[4] ENCKE, *Ueber d. Methode d. kleinsten Quadrate*, Berliner Astr. Jahrb., 1834 249.
CHAUVENET, Manual of Spherical and Practical Astronomy, 469, Phila., 1864.
HELMERT, Die Ausgleichungsrechnung, Leipzig 1872.
MEYER, Vorlesungen ü. Wahrscheinlichkeitsrechnung, 243, Leipzig 1879.
MERRIMAN, Method of Least Squares, New York 1894.
WEINSTEIN, Handbuch der physikal. Maassbestimmungen, I 54, Berlin 1886.

3. *Algebraic selection.* Means can be classed according to the way in which they are computed.

a. Combinatory means.[1] These are

$$F_1(x)=\sqrt[1]{\frac{\Sigma x}{n}}$$

$$F_2(x)=\sqrt[2]{\frac{1.2.\ \Sigma_{\alpha,\beta}\ x_\alpha x_\beta}{n(n-1)}}$$

$$\vdots \qquad \vdots$$

$$F_r(x)=\sqrt[r]{\frac{1.2.\ .\ .\ .\ r.\ \Sigma_{\alpha,\beta,\ .\ .\ .,\rho}\ x_\alpha x_\beta\ .\ .\ .\ x_\rho}{n(n-1)\ .\ .\ .\ (n-r)}}$$

$$\vdots \qquad \vdots$$

$$F_n(x)=\sqrt[n]{x_1\,x_2\ .\ .\ .\ x_n}.$$

Of these means only two have come into use, namely, the arithmetic mean

$$F_1(x)=\frac{x_1+x_2+\ .\ .\ .\ +x_n}{n}=A.$$

and the geometric mean

$$F_n(n)=\sqrt[n]{x_1 x_2\ .\ .\ .\ x_n}=G.$$

b. Power means. These are

$$p_1(x)=\sqrt[1]{\frac{\Sigma x^1}{n}}$$

$$p_2(x)=\sqrt[2]{\frac{\Sigma x^2}{n}}$$

$$\vdots \qquad \vdots$$

$$p_n(x)=\sqrt[n]{\frac{\Sigma x^n}{n}}.$$

---

[1] SCHEIBNER, *Ueber Mittelwerthe,* Ber. d. math.-phys. Cl. d. könig̣. sächs. Ges. d. Wiss., 1873 XXV 562.

FECHNER, *Ueber d. Ausgangswerth d. kl. Abweichungssumme,* Abhandl. d. math.-phys. Cl. d. könig̣. sächs. Ges. d. Wiss., 1878 XI 1(76).

By analogy we might put

$$p_0(x) = \sqrt[0]{\frac{\Sigma x^0}{n}};$$

but since $x^0 = 1$, then $p_0(x) = \sqrt[0]{1} = 1^{\frac{1}{0}} = 1^\infty$, which represents an indeterminate quantity[1] between 0 and ∞. This can be taken to express the fact that the median $M = p_0(x)$ is not dependent on the numerical values of $x_1, x_2, \ldots, x_n$.

The value

$$p_1(x) = \frac{x_1 + x_2 + \quad . \quad . \quad . \quad + x_n}{n} = A$$

is the arithmetic mean. When x represents the errors from an average, their mean

$$p_2(x) = \sqrt[2]{\frac{x_1^2 + x_2^2 + \quad . \quad . \quad . \quad + x_n^n}{n}} = m$$

is the mean-square-error as used in the method of least squares. The quartic mean $p_4(x)$ has also been used in the calculation of precision.[2]

4. *Selection to minimize a function of the variations.* Means have been so selected as to make

$$\frac{d(f(v))}{dR} = 0,$$

where

$$v_k = x_k - R, \qquad (k = 1, 2, \ldots, n).$$

For $f(v) = \Sigma(v)$, we have[3]

---

[1] Cauchy, Cours d'analyse algebr., 69, Paris 1821.

[2] Gauss, Theoria combinationis observationum, I, 11.

[3] Boscovich, De littera expeditione ad dimetiendos duos meridiani gradus, Romae 1755. (I have not seen this work. It is described in Todhunter, Hist. Theories Attract., I 305, 332, Lond. 1873.)

Boscovich, *De recentissimus graduum dimensionibus*, Philosophia recentior a B. Stay, II 420, Romae 1760. (I have not seen this work. It is described in Todhunter, Hist. Theorics Attract., I 321, Lond. 1873.)

Laplace, *Mém. sur la prob.*, Mém. de math. et de phys. par divers savans, Acad. Paris, 1774 VI 621 (635).

Bernoulli, *Milieu*, Encyclopédie méthodique, Math., II 404, Paris 1785: Dict. encycl. d. math., Paris 1789.

Laplace, *Sur les degrés mesurés des méridiens*, Hist. Acad. Sci. Paris 1789, Mém. de math., 18.

Laplace, Mécanique celeste, III 40, Paris 1804; Oeuvres, II 144.

Laplace, Théorie anal. des prob., Suppl. 2, §2.

Fechner, *Ueber d. Ausgangswerth d. kl. Abweichungssumme*, Abhandl. d. math.-phys. Cl. d. königˇ sächs. Ges. d. Wiss, 1878 XI 1.

Glaisher, *On the law of facility of errors of observations*, Mem. Roy. Astr. Soc. Lond., 1872 XXXIX 75 (123).

$$\frac{d\Sigma(v)}{dR}=0,$$

and

$$R_0=M.$$

For $f(v)=\Sigma v^2$, or, what amounts to the same thing, for

$$f(v)=\sqrt[2]{\frac{\Sigma v^2}{n}}=m,$$

we have[1]

$$R_1=\frac{x_1+x_2+\quad.\quad.\quad.\quad+x_n}{n}=A$$

Average.

Average and centroid. The average is defined as

$$\mathrm{A}=\frac{x_1+x_2+\quad.\quad.\quad.\quad+x_n}{n}.$$

If m_1, m_2, . . ., m_r denote the number of times the values x_1, x_2, . . ., x_r occur respectively, then

$$\mathrm{A}=\frac{x_1+x_2+\quad.\quad.\quad.\quad+x_n}{n}=\frac{m_1\,x_1+m_2\,x_2+\quad.\quad.\quad.\quad+m_r\,x_r}{m_1+m_2+\quad.\quad.\quad.\quad+m_r}$$
$$=\frac{\Sigma mx}{\Sigma m}.$$

The average thus corresponds to the centroid of a system of parallel forces.

If the results are so numerous that the values of x can be treated as continuous and $\phi(x)dx$ can be substituted for m, the average can be substituted, with a small error ϵ, for the abscissa of the centre of gravity. Thus

$$\mathrm{A}=\frac{\int_{x_1}^{x_n}x\phi(x)dx}{\int_{x_1}^{x_n}\phi(x)dx}\pm\epsilon,$$

which in consideration of (2) and (4) becomes

---

[1] Legendre, Nouv. méthodes pour la détermination des orbites des comètes, VIII, Paris 1805.

Merriman, *List of writings relating to the method of least squares*, Trans. Conn. Acad., 1877 IV 151.

$$A = \int_{x_1}^{x_n} x\phi(x)dx \pm \epsilon.$$

If $\phi(x)$ is unknown and the number of results is large, the average A represents the centroid-abscissa x_ξ with a degree of certainty and within limits determined by POISSON.[1] For a large number of results

$$A = x_\xi \pm \frac{2\gamma\sqrt{h}}{\sqrt{n}}, \tag{12}$$

with a probability of

$$\Phi(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2} dt.$$

where h is a quantity derived from the means of the first and second powers of the errors but is not amenable to practical calculation except for known $\phi(x)$.

Various other considerations bearing on the relation of the average to the centroid and to the individual results,[2] although necessary to a just appreciation of these relations, cannot be touched here. The assertion that the use of the average is the mean supposition of all possible suppositions as to the mode of obtaining value,[3] in addition to its questionable character,[4] rests on the assumption of symmetrical probability which cannot be admitted here.

Precision of the average. The calculation of the mean variation, the mean-square-error, the probable error, the constant of precision, etc., are to be found in the numerous works on measurement. They generally start with the assumption of a symmetrical curve of probability and pass over assymmetrical curves as being symmetri-

---

[1] POISSON, Recherches sur la probabilité des jugements, ch. IV, Paris 1837.

[2] LAGRANGE, *Mém. sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations*, Miscell. Taurinensia, 1770–1773, V (math.) 167; Oeuvres, II 171.

ENCKE, *Ueber d. Anwendung d. Wahrscheinlichkeits-Rechnung auf Beobachtungen*, Berliner Astr. Jahrb. f. 1853, 310.

PIZZETTI, *Sopra una generalizzazione del principo della media aritmetica*, Atti d. R. Accad. dei Lincei, Rend., 1889 (4) V_1 186.

[3] DE MORGAN, *On the theory of errors of observation*, Trans. Camb. Phil. Soc., X 409 (416).

[4] GLAISHER, *On the law of facility of errors of observation*, Mem. Roy. Astr. Soc. Lond., 1872 XXXIX 75 (90).

cal curves with systematic errors. An elementary treatment without this assumption is given by BERTRAND.[1]

Numerical error of the average. The computation of the average involves a decision on the number of decimal places to be retained in the observed values and in the mean after division.

The limitation of the number of decimal places used in writing a result introduces an error into each result. Each result differs from the truth by not more than one half-unit in the last place; its mean error is $\frac{1}{4}$ of a unit in the last place, or 0.25β where β denotes 1 unit in the last place.

When n results are added to form an average, the mean error of the sum will be $0.25\beta\sqrt{n}$. The average itself will have a mean error of $\frac{0.25}{\sqrt{n}}\beta$.

Although this supposition is the usual one for numerical work, it is not strictly true. Such a treatment of the mean error is valid only when the law of frequency is expressed by (11). The error thus introduced is, however, negligible.

Since it is understood that the last decimal place is subject to a mean error of 0.25β, the extra decimals obtained in calculating the average of 100 results may be retained to one place beyond the original results.

Even when the number of results is less than 100, the retention of one place further introduces less error than rounding off to a places. Thus for 25 results the mean error of the a place is 0.05β, and of the $a+1$ place is $0.50\beta'$. If we round off to the a place on a supposition of a mean error of 0.25β in the usual way, the uncertainty of the a place is increased.

Given the set of 9 values 213, 215, 213, 210, 212, 214, 215, 210, 212. The mean numerical error of each value is 0.25β. The mean error of the average will be

$$\frac{0.25\beta}{\sqrt{9}}=0.08\beta, \text{ or } 0.83\beta'.$$

The average is

$$\frac{1914}{9}=212.666\ldots\ldots$$

which is subject to a mean error of 0.08β in the first place before the point or $0.8\beta'$ in the first decimal place. If we round off to 213,

---

[1] BERTRAND, Calcul des probabilités, ch. X, Paris 1889.

we introduce an uncertainty corresponding to a mean error of 0.25β in the whole numbers, whereas by retaining the 6, the uncertainty corresponds to only 0.08β in the unit-place or to $0.8\beta'$ in the first decimal place, being a gain corresponding to 0.17β. Rounding off to 212.7 adds an uncertainty of $0.25\beta'$ in the first decimal place, giving a total of $0.83\beta'+0.25\beta'=1.08\beta'$ for that place or 0.108β for unit-place, being a loss corresponding to 0.028. Since under any circumstances the loss would have to be 0.025β, the writing of 212.67 has practically no advantage over 212.7.

The case is different when the uncertainty of the values is not due simply to the omission of decimal places. Let the measure of the uncertainty of a value x be denoted by $\pm\,\Delta x$. Then the uncertainty of the average of n results will be given by

$$\frac{\pm\,\Delta x}{\sqrt{n}}.$$

The mean error is the most convenient measure of uncertainty.

Under very favorable conditions the record of the Hipp chronoscope[1] is liable to a mean error of $\Delta x=1.5^{\sigma}$. The average of 9 records is reliable to 0.5^{σ}. To obtain a result numerically precise to 1^{σ}, i. e. with a mean error of 0.25^{σ}, it would be necessary to have 36 original records.

Thus, if a body, e. g. a control-hammer, were known to fall with perfect constancy, 36 records with the chronoscrope would be required to determine its time of fall to 1^{σ}.

Dependence on characteristic variations. The result of a set of direct measurements is stated to be $A\pm d$, $A\pm m$ or $A\pm r$. The quantity d is the mean variation or mean error, m is the mean-square-error and r is the half probable variation or probable error. These characteristic variations are determined from the formulas

$$d=\frac{(v_1)+(v_2)+\ .\ \ .\ \ .\ +(v_n)}{\sqrt{n(n-1)}}\left(1\pm\frac{0.756}{\sqrt{n-1}}\right)$$

$$m=\sqrt{\frac{v_1^2+v_2^2+\ .\ \ .\ \ .\ +v_n^2}{n-1}}\left(1\pm\frac{0.708}{\sqrt{n-1}}\right)$$

$$r=0.674\sqrt{\frac{v_1^2+v_2^2+\ .\ \ .\ \ .\ +v_n^2}{n-1}}\left(1\pm\frac{0.708}{\sqrt{n-1}}\right)$$

In these formulas the signs $\pm$ have the meanings usually given them in works on adjustment.

---

[1] Külpe and Kirschmann, *Ein neuer Apparat zum Controle zeitmessender Instrumente*, Phil. Stud., 1892 VII 145.

These results indicate the range within which any other single measurement under the same circumstances may be expected to differ from A with a given degree of probability. Thus, we can wager 1 to 1 that a repetition of the measurement under the same conditions will give a result differing from A by not more than r.

In the final statement of the average it is unnecessary and misleading to use more figures than would be justified by the characteristic variations.

Thus if a set of 25 measurements on reaction-time gives an average of 0.2346 sec. with a probable error of 0.012 sec., the second figure, 3, of the result is uncertain to the extent of more than ± 1 unit. The third figure, 4, of the result is uncertain to the extent of ± 12 units, and the last figure, 6, to the extent of ± 120 units. As figures when rounded-off are understood to be uncertain to the extent of a mean error of ± 0.25 unit in the last place, the statement that the result is 0.23 sec. is somewhat less reliable than the figures themselves indicate and the statement that the result is 0.235 or 0.2346 is quite misleading. The usual method, whereby the average is given to the last place justified by the compution, while the amount of d, m or r is independently stated, is justifiable or not according to the purpose in hand. When the purpose is simply the determination of an average, there can be no ground for affixing meaningless decimal places; the average should not be stated further than the first place rendered insecure by the characteristic variation.

Median.

Median and middle value. The median is that value which is obtained by counting off an equal number from each end of the series of results arranged singly according to size. If $p_1, p_2, \ldots, p_n$ represent the relative frequencies of $x_1, x_2, \ldots, x_n$, then

$$\sum_{x_1}^{M} p = \sum_{M}^{x_n} p.$$

The values $x_1, x_2, \ldots, x_n$ have all an equal influence in the determination of x_m. Each quantity x_k is either above or below x_m; how far above is not regarded. Those above x_m might be called positive results and those below x_m might be called negative. Thus we might put

$$\sum_{x_m}^{x_n} p = r \qquad \sum_{x_1}^{x_m} p = s$$

and

$$r + s = \mu$$

or

$$r = s = \frac{\mu}{2}.$$

The relation between M and x_m can be determined by BERNOULLI'S theorem. When μ is large,

$$M = x_m \pm \gamma r \sqrt{\frac{2}{\mu}} \tag{13}$$

with a probability of

$$P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2}\, dt + \frac{\sqrt{\mu} \,.\, e^{-\gamma^2}}{r\sqrt{2\pi}}$$

The values of γ are to be determined from the usual table for

$$\Phi(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2}\, dt.$$

Computation of the median. The median is defined as that value which occupies the position given by

$$\frac{x_1^0 + x_2^0 + \;.\;\;.\;\;.\; + x_n^0 + 1}{2} = \frac{n+1}{2}$$

in the series of values $x_1, x_2, \ldots, x_n$ taken in order of size from the smallest to the largest and from the largest to the smallest.

Let the number of occurrences of each value of x be denoted by $m_a, m_b, \ldots, m_r$. The series of differing values $x_1, x_2, \ldots, x_r$, finitely expressed, can be regarded as having arisen from the series $x_1, x_2, \ldots, x_n$ expressed each to an infinite number of decimals by rounding-off all the decimals to the a place. In the a place the set $x_1, x_2, \ldots, x_a$ will all agree and can be expressed by $m_a\, x_a$. Likewise we have the sets $m_b\, x_b, \ldots, m_r\, x_r$.

When these sets are arranged in order of size

$$m_a\, x_a,\; m_b\, x_b,\; \ldots,\; m_{l-1}\, x_{l-1},\; m_l\, x_l,\; m_{l+1}\, x_{l+1},\; \ldots,\; m_r\, x_r$$

the set containing the median will be $m_l\, x_l$ where

$$(m_a + m_b + \;.\;\;.\;\;.\; + m_{l-1}) - (m_{l+1} + m_{l+2} + \;.\;\;.\;\;.\; + m_r) < m_l.$$

The set containing the median is not necessarily the middle set.

The median will be one of the m_l values which have been rounded

off to the same value x_l. Each value x_l represents some value $x_l \pm 0.5\beta$ where β indicates 1 unit of the order a.

When

$$(m_a+m_b+\ .\ \ .\ \ .\ +m_{l-1})-(m_{l+1}+m_{l+2}+\ .\ \ .\ \ .\ +m_r)=-c,$$

the median value occupies a place among the m_l values given by c where

$$(m_a+m_b+\ .\ \ .\ \ .\ +m_{l-1})+c+\frac{m_l-c}{2}=\frac{m_l-c}{2}+(m_{l+1}+m_{l+2}+\ .\ \ .\ \ .\ +m_r).$$

The median is thus not the middle value of the group m_l but of the group m_l-c.

If the whole interval from which the values of x_l were derived be denoted by S, then the position of M within the interval S will be given by

$$M=x_l+\frac{c}{2m_l}S$$

If the extreme values for x_l be

$$x_l'=x_l-0.5\beta$$
$$x_l''=x_l+0.5\beta$$

where β is 1 unit of the last place, then

$$M=x_l+\frac{c}{2m_l}\beta$$

For the results given on p. 17, 210 is the smallest value, 210 the next, 212 the next, etc. As there are 9 values, the median will be the $\left(\frac{9+1}{2}=5\right)$th value from the smallest; this is 213. The largest value is 215, the next 215, the next 214, etc. The 5th from the largest is 213.

The 4th value from the largest is also 213^σ. Thus $x_l=213^\sigma$, $m_l=2$, $c=+1$ and $\beta=1^\sigma$. Consequently

$$M=213^\sigma+\tfrac{1}{4}1^\sigma=213.25^\sigma.$$

The following example of calculating the median is given by Fechner.[1] As a historical interest is present, the rather naïve method of using decimals is left untouched.

"Take the case where the result runs as follows :

result	1	2	3	4	5
number of times it has occurred	2	5	16	10	7

[1] Fechner, *Ueber d. Ausgangswerth d. kleinsten Abweichungssumme*, Abhandl. d. math.-phys. Cl. d. k. sächs. Ges. d. Wiss., 1878 XI 1 (19).

"The total number n is here 40, and the $\frac{n+1}{2}$., i. e. the $20\frac{1}{2}$. value counted from either left or right end of the series cuts into the number 16; thus the median is to be sought by interpolation in the series of the 16 results giving the value 3, but as the 20th, not as the $20\frac{1}{2}$th. The limits of this series of 3s are 2.5 and 3.5. Counting from the left side we have 7 values up to the limit 2.5 of the series and 13 more are needed to make the 20th value which falls among the 3s; thus according to the simplest principle of interpolation we have to take from the limit 2.5 upward still $\frac{13}{16}=0.8125$ of this interval in order to reach M, making $M=2.5+0.8125=3.3125$. Going from the right hand end, we have 17 values up to the limit 3.5 of the series, and lack still 3 of the 20, which fall in the series. Thus we have to subtract from the limit $\frac{3}{16}=0.1875$ of the interval in order to reach M, which gives $3.5-0.1875=3.3125$ as before."

Numerical error of the median. The mean numerical error of a single value being 0.25β, that for m_l values will be

$$\frac{0.25}{\sqrt{m_l}}\beta.$$

In the foregoing example where $m_l=2$, the mean error of the result 213^σ is 0.18β and of the extended result 213.3^σ is $1.8\beta'$, where $\beta=1^\sigma$ and $\beta'=0.1^\sigma$.

As a slightly different example take the values 44, 51, 46, 50, 47, 49, 47, 45, 48, 50. The median will be the $\left(\frac{10+1}{2}\right)$th or $5\frac{1}{2}$th value. The fifth from the smallest is 47; the fifth from the largest is 48; the $5\frac{1}{2}$th will lie between the two. As there is no reason to prefer one extreme of the interval $48-47$ to the other, the simplest method is to take $M=47.5$. The numerical uncertainty of 48 is represented by a mean error of 0.25β; that of 47, derived from two results by $\frac{0.25\beta}{\sqrt{2}}=0.18\beta$. The mean error of their sum will be

$$\sqrt{(0.25)^2+(0.18)^2}\beta=0.31\beta\,;$$

and of their average $\qquad \frac{0.31\beta}{2}=0.16\beta.$

The numerical mean error for the median in this example is thus 0.16β for the unit-place.

The decimal place, already uncertain by a mean error of $0.25\beta'$, becomes uncertain to the extent of a mean error of $1.85\beta'$. Although 47.5 is less uncertain than 47 or 48, it has nevertheless quite a numerical uncertainty.

The numerical uncertainty of the average of the ten results would be indicated by a mean error of $\frac{0.25\beta}{\sqrt{10}}=0.08\beta$ for the unit-place or $0.8\beta'$ for the first decimal. The median is thus at a disadvantage numerically.

As the results gather more around a middle value in more accurate work, more values will coincide with the median, m_l will become larger and the numerical error for a given number of results will become less. The numerical error of the average will remain the same.

As mentioned on pages 2 and 17 the influence of $\phi(x)$ on the numerical error is negligible. Thus in calculating the numerical error of the median, as long as the unit of number does not exceed the size of the mean error, we can safely suppose the values x_l to have arisen by rounding-off equally frequent values throughout β.

Dependence of accuracy on the number of results. Gauss has deduced an expression for the accuracy of the mean value of any power of a variation as depending on the number of variations.[1] With slight changes the results can be stated as follows on the supposition of (11). Let x_κ be the mean as determined from the κ powers of the n observations. Then with not too small numbers of results the probable uncertainty, in the same sense as the probable error, for $x_0=M$ determined from $x_1^0, x_2^0, \ldots, x_n$ is $\pm\frac{0.752}{\sqrt{n-1}}$; for $x_1=A$ determined from $x_1^1, x_2^1, \ldots, x_n^1$ it is $\pm\frac{0.510}{\sqrt{n-1}}$.

Other things being equal, it is necessary to take 249 observations to gain the same accuracy for the median as is given by 114 observations for the average.

Dependence on characteristic variations. As noted on p. ? the median is that representative value which corresponds to a minimum for the sum of the absolute values of the first powers of the variations. The mean variation from the median will bear to the median a relation similar to that which the mean-square-error bears to the arithmetic mean. The median will thus be stated as

$$M\pm a,\ M\pm l \text{ or } M\pm s, \text{ where}$$

---

[1] Gauss, *Bestimmung d. Genauigkeit d. Beobachtungen*, Zt. f. Astr., 1816 I 185; Werke, IV 109.

Lipschitz, *Sur la combinaison des observations*, C. R. Acad. Paris, 1890 CXI 163.

$$a = \frac{\Sigma\,(x-M)}{\sqrt{n(n-1)}}\left(1 \pm \frac{0.756}{\sqrt{n-1}}\right).$$

$$l = \sqrt{\frac{\Sigma\,(x-M)^2}{n-1}}\left(1 \pm \frac{0.708}{\sqrt{n-1}}\right)$$

$$s = 0.674\ l.$$

The number of significant figures to be retained in stating the median is regulated in the same way as for the average. For the example given on p. 17, the median is calculated to be 213.3^σ; the mean variation is 1.6^σ. The last figure is uncertain by at least 1.6^σ and cannot justifiably be used for final statement. Even the last figure in 213^σ is slightly uncertain.

In FECHNER's example on p. 21 the mean variation is 1.14. This makes even the whole number 3 rather uncertain and for final statement renders utterly valueless the four-place decimal.

In the example given on p. 22, the mean variation of the individual results from the median is 2.0. Whence it follows that for the mere statement of a representative result, the decimal place is totally worthless and even the unit-place is unreliable. The mean variation from the average 47.7 is also 2.0. The conclusion is the same. Thus the average in a case of this kind has no advantage whatever.

As the mean variation becomes less, owing to better agreement of the results, the numerical mean error of the median will also decrease, whereas that of the average will remain the same. Thus the numerical advantage of the average is valueless, for a final statement, when the mean variation is large, and this advantage itself is lost as the mean variation decreases.

The labor of obtaining the extra decimal is thus not justified when the results disagree to such an extent. By a preliminary estimate or computation or by a cursory glance at the values themselves it is generally possible to determine the number of places required and thus to adjust the amount of the labor to the worth of the result.

In the consideration of the characteristic variations both for the median and for the average, I have, for the sake of using GAUSS's deductions, made the supposition that the law of error is

$$\phi(v) = \frac{h}{\sqrt{\pi}} e^{-h^2 v^9}. \qquad (11)$$

Labor of computation. There is one important property of the median which can be understood only after practical acquaintance with it, namely, its economy. Suppose the original results of the

example already used to be set down in a line or a column: 213, 215, 213, 210, 212, 214, 215, 210, 212. Let the higher numbers be marked off successively by a small check, thus 2ĭ5, till 5 have been checked. This is the median. After a little practice this can be done for 10 or 25 values with unexpected rapidity. The result is rapidly verified by checking off the numbers from the smallest upward. The determination of the average requires the addition of 9 figures and the division of the result by 9. Even in this example where the first two figures can be neglected and the whole work can be done mentally, yet the time required is much longer.

In all examples mentioned, owing to the size of the mean variation there would have been no gain in computing the average, whereas the additional labor would have been a decided loss. When it is remembered that saving in labor in computation means additional opportunity for obtaining results, it is justifiable to claim that the most economical method of computation should be employed.

In the typical examples given the additional uncertainty of the median is negligible in comparison with the characteristic variations. When this is not the case, it is a question whether to obtain double the number of results for the median or to perform the additional computation required by the average, in order to obtain the mean with a given precision.

Median error. On the assumption of (11) the probable error should be nearly the same whether determined by Bessel's formula

$$r_1 = 0.674 \sqrt{\frac{v_1^2 + v_2^2 + \quad . \quad . \quad . + v_n^2}{n-1}}$$

or by Peters's formula

$$r_2 = 0.845 \; \frac{(v_1) + (v_2) + \quad . \quad . \quad . + (v_n)}{\sqrt{n\,(n-1)}}$$

or by counting off in order of size till the middle error is reached. The probable error is thus in the last case the median error. In the same way that the mean error for the median corresponds to the mean-square-error for the average, so the median error for the median would correspond to the mean error for the average. In a similar manner the median error for the median can be compared with the probable error for the median, in order to test the validity of the assumption mentioned.

Other discussions on the median. In addition to the productions cited elsewhere, several other articles containing accounts of or refer-

ences to the median have been consulted. Those whose titles I have noted down are by COURNOT,[1] EDGEWORTH,[2] TURNER,[3] SCRIPTURE,[4] VENN.[5]

WEIGHT AND INFLUENCE.

In forming a direct average each measurement in a given set of n measurements has an influence of

$$f_1=\frac{x_1}{n}\ ;\ f_2=\frac{x_2}{n}\ ;\ .\ \ .\ \ .\ ;\ f_n=\frac{x_n}{n}$$

on the result. The influence of a quantity is thus directly proportional to its numerical value. The numerical values x_1, x_2, . . ., x_n can thus be called the relative influences of the 1, 2, . . ., nth measurements.

In combining results from different sets of measurements it is not always desirable for them to influence the average in direct proportion to their face-values. This has led to the use of a system of multipliers, called weights, by which the influence of a quantity is modified. The various results are each multiplied by a coefficient $p_1, p_2,$ . . ., p_n such that $\Sigma p=n$. The average of the weighted results will be the "weighted mean,"

$$A_p=\frac{p_1x_1+p_2x_2+\ .\ \ .\ \ .\ +p_nx_n}{p_1+p_2+\ .\ \ .\ \ .\ +p_n} \tag{14}$$

The fact that in concrete cases Σp is not equal to n arises from a tacit division of both numerator in (14) by the same number.

The weighted mean agrees with the direct mean only when

$$p_1=p_2=\ .\ \ .\ \ .\ =p_n. \tag{15}$$

It is a very natural step to apply this concept to individual measurements; some measurements are naturally better than others. But when all measurements have been made apparently with equal care and there is no reason to prefer one to another, it may be said that there is no *a priori* reason for weighting one different from

[1] COURNOT, Exposition de la théorie des chances et des probabilités, 120, Paris 1843.

[2] EDGEWORTH, *On discordant observations*, Phil. Mag., 1887 (5) XXIII 364.

EDGEWORTH, *New method of reducing observations*, Phil. Mag., 1887 (5) XXIV 222.

EDGEWORTH, *Empirical proof of the law of error*, Phil Mag, 1887 (5) XXIV 330.

EDGEWORTH, *Discussion on Dr.* VENN'S *paper*, Jour. Roy. Statist. Soc. Lond., 1891 LIV 453.

[3] TURNER, *On Mr.* EDGEWORTH'S *method of reducing observations*, Phil. Mag., 1887 (5) XXIV 466.

[4] SCRIPTURE, *On the adjustment of simple psychological measurements*, Psych. Rev., 1894 I 281.

[5] VENN, *On averages*, Jour. Roy. Statist. Soc. Lond, 1891 LIV 429.

another. The simplest assumption is the purely arbitrary one that all are of equal weight.

This gives a very good result if the values of x run along very regularly and close together. No dissatisfaction is felt as long as the individual variations fall within the limits.

$$-l < v_k < +l$$

or[1]

$$v_k < (l)$$

where

$$v_k = x_k - A, \quad (k = 1, 2, \ldots, n), \tag{16}$$

l not being considered a large quantity. But if some very large value x_r occurs so that

$$(v_r) > (l),$$

the natural supposition is that x_r is not so reliable as the other values of x. Sometimes it is rejected, i. e. the weight $p_r = 0$ is attached to it while $p_1 = p_2 = \ldots = p_{r-1} = p_{r+1} = \ldots = p_n = 1$. Sometimes it receives a weight $p_r = \frac{1}{2}$, or $p_r = \frac{1}{3}$, while the others receive $p = 1$, in a purely arbitrary fashion.

To know when to reject values it is necessary to assign some value to l. This has led to various criteria for rejection, the best known of which are those of PEIRCE,[2] CHAUVENET[3] and STONE.[4]

---

[1] (x) is used for abs x, or x taken without regard to sign.

[2] PEIRCE, *Criterion for the rejection of doubtful observations*, Astr. Jour. (Gould), 1852 II 161.

GOULD, *Report . . . containing directions and tables for the use of* PEIRCE'S *criterion*, U. S. Coast Surv., Rept. 1854. 131*.

GOULD, *On* PEIRCE'S *criterion for the rejection of doubtful observations, with tables*, Astr. Jour. (Gould), 1855 IV 81.

AIRY, *Letter from . . .* [remarks on PEIRCE'S criterion], Astr. Jour. (Gould), 1856 IV 137.

WINLOCK, *On Prof.* AIRY'S *objection to* PEIRCE'S *criterion*, Astr. Jour. (Gould), 1856 IV 145.

NEWCOMB, *A generalized theory of the combination of observations so as to obtain the best results*, Am. Jour. Math., 1886 VIII 343. (The note on p. 344 contains two striking deductions.)

[3] CHAUVENET, Manual of Spherical and Practical Astronomy, 564, Phila. 1864.

[4] STONE, *On the rejection of discordant observations*, Month. Not. Roy. Astr. Soc. Lond., 1868 XXVIII 165.

GLAISHER, *On the rejection of discordant observations*, Month. Not. Roy. Astr. Soc. Lond., 1873 XXXIII 391.

STONE, *On the rejection of discordant observations*, Month. Not. Roy. Astr. Soc. Lond., 1873 XXXIV 9.

GLAISHER, *Note on a paper by Mr.* STONE . . . , Month. Not. Roy. Astr. Soc. Lond., 1874 XXXIV 251.

STONE, *Note on a discussion . . .* , Month. Not. Roy. Astr. Soc. Lond., 1875 XXXV 107.

The criteria for rejection have never proven satisfactory. The general sentiment seems to be that the rejection of an honestly made observation simply because it differs largely from the expected value amounts to an attempt to make the work appear more accurate than it is.[1] The rejection of observations by calculation of the average first with it and then without it, is said to be like what happens in war when two detachments of the same army meet in the dark and fire into each other, each supposing the other to belong to the common enemy.[2]

The very questionable justification for any rejection of results on account of their divergence has led to various systems of weights.[3] The main objections to these systems are 1. the assumption of the validity of SIMPSON's law[4] $\phi(-v_k)=\phi(+v_k)$, where v_k is defined as in (16) and ϕ indicates the relative number of times of occurrence of v_k; 2. the labor of computation which is often out of all proportion to the gain.[5]

The rejection of observations has been a troublesome question for psychologists. The results did not agree and absolutely refused to group themselves around the arithmetical mean. It was not a question of a single result differing from all the rest but of several results tending toward an extreme value. This led to a process of wholesale rejection[6] which reached a crisis in giving double sets of results, once

---

[1] HALL, *Orbit of Iapetus*, 40, Astr. and Meteor. Obs. for 1882, U. S. Naval Obs., App. I., Wash. 1885.

NEWCOMB, *A generalized theory of the combination of observations so as to obtain the best result*, Am. Jour. Math., 1886 VIII 343 (345).

FAYE, *Sur certains points de la théorie des erreurs accidentelles*, C. R. Acad. Sci. Paris, 1888 CVI 783.

[2] DOOLITTLE, *The rejection of doubtful observations*, Wash. Bull. Philos. Soc., 1884 VI 152, in Smithsonian Misc. Coll., 1888 XXXIII.

[3] DE MORGAN, *Theory of probability*, Encyc. Metropol., II 456, Lond. 1847.

GLAISHER, *On the law of facility of errors of observation and on the method of least squares*, Mem. Roy. Astr. Soc. Lond., 1871 XXXIX Pt. I. 75 (103).

NEWCOMB, *A generalized theory of the combination of observations so as to obtain the best result*, Am. Jour. Math., 1886 VIII 343.

SMITH, *True average of observations*, Nature, 1888 XXXVII 464.

[4] SIMPSON, *An attempt to show the advantage arising by taking the mean*, Misc. Tracts, 64, Lond. 1757.

[5] EDGEWORTH, *Choice of means*, Phil. Mag., 1887 (5) XXIV 268 (271).

[6] EXNER, *Exper. Untersuch. d. einfachsten psych. Processe*, Arch. f. d. ges. Physiol. (Pflüger), 1873 VII 601 (613).

V. KRIES and AUERBACH, *Die Zeitdauer einfachster psychischer Vorgänge*, Arch. f. Physiol. (Du Bois-Reymond), 1877 297 (307).

without rejection and then with rejection,[1] and found its *reductio ad adsurdum* in making experiments in sets of 25 of which 20 were selected for the calculation of the average.[2]

All this has made it evident that, even if the arithmetic mean of the observations is justifiable in the physical sciences, it is not *a priori* the best representative value for psychological measurements.

The trouble lies in the fact that, because the average is the most plausible representative for certain kinds of differing quantities, it has been treated as the best one in all cases.[3] In the use of the average each individual quantity influences the result in direct proportion to its numerical value. The value $x_i = a$ contributes to the result $\frac{a}{b}$ times as much as the value $x_k = b$. This is unquestionably the correct method to pursue when the individuality of the quantity is of no account. If r cubic meters of soil must be removed for a railroad-cut, it makes little difference just how much each particular car of a train carries, provided the average is satisfactory. Each car counts not merely as an overloaded or an underloaded car, but as overloaded or underloaded to a definite extent.

This is not the case in most measurements. The measurements tend to group themselves around some mean value, and when a widely different value occurs it is looked upon with distrust. To use it in an average is to make it count, not as one single value above or below the mean, but as an individual counting for every unit of divergence. The more it differs, the more it counts.[4] For example, in a set of values 3, 4, 2, 2, 3, 10, it is evident from mere inspection that the grouping is around 3. The average is 4, because the one extreme value 10 contributed to the formation of the average as much as the four values 3, 2, 2, 3 put together. If we had 15 instead of 10, the influence of this extreme value would have been more than that of all the rest. In the particular case of measurements the more a value differs from the rest, the less we think of it; if it

[1] BERGER, *Ueber d. Einfluss der Reizstärke auf d. Dauer einfacher psych. Vorgänge*, Phil. Stud., 1886 III 38 (61).

CATTELL, *Psychometrische Untersuchungen*, Phil. Stud., 1886 III 305 (317).

[2] JASTROW, *Studies from the Univ. of Wisconsin*. Am. Jour. Psych., 1892 IV 382 (413).

[3] DE MORGAN, *On the theory of errors of observation*, Trans. Camb. Phil. Soc., X 409 (416).

VENN, *On averages*, Jour. Roy. Statist. Soc. Lond., 1891 LIV 429.

[4] BOWDITCH, Note to LAPLACE'S Mécanique celeste, translated, vol. II, 434, Boston 1832.

differs too much, we think so little of it that we are tempted to throw it out altogether.

Instead of allowing the individual result to have an influence proportional to its value, why not let it enter into the formation of the mean as one individual differing from the mean regardless of the amount of the difference?

Thus, if the graduating class in college happens to contain one very tall man, the average height will be greater than the averages for other years. The tall man contributes to the representative value not simply as one man but with the influence of several men.

It would seem more natural that each individual should influence the representative value merely as one individual. If all the men were placed in order of height, the most natural representative would be the man in the middle, or the man who would be determined by counting off an equal number of individuals from both extremes. This is equivalent to determining M by the formula on p. 20; and the height of this middle man is the median height. If there happens to be a very tall man among them, a few millimeters more or less in his height will make a difference in the average, but as long as he remains taller than the middle man the median will remain the same.

Let the set of results be indicated by x_1^0, x_2^0, . . ., x_n^0, where these letters are used merely as symbols for single quantities. Thus, if the measurements are the heights of a set of individuals, x_1^0 will designate a certain individual; if they are observations, each will designate an observation. In taking the average, we influence each individual quantity with a number indicating its numerical value. Thus

$$\frac{\Sigma x x^0}{n} = \frac{x_1 x_1^0 + x_2 x_2^0 + \ . \ \ . \ \ . + x_n x_n^0}{n} = A.$$

But let a tall man count as only one tall man regardless of just how tall he is, provided he is above the mean, and let a short man count likewise as one man. Since tall and short are only relative terms, there must be some value above which all the men are to be called tall and below which they are to be called short. This is the value that we have called the middle value. Each individual counts as one unit. In general terms,

$$\frac{x_1^0 + x_2^0 + \ . \ \ . \ \ . + x_n^0}{n} = \frac{\Sigma x^0}{n} = M^0.$$

Each value thus has the same influence.

The influence of an individual measurement can be defined as its relative effect in the formation of the mean; its weight is an arbi-

trary multiplier prefixed to its numerical value. The influence of a measurement in taking an average is thus the product of its numerical value by its weight.

The general formula (14) for the weighted mean becomes in the notation just used,

$$A = \frac{p_1 x_1 x_1^0 + p_2 x_2 x_2^0 + \ . \ . \ . \ + p_n x_n x_n^0}{p_1 + p_2 + \ . \ . \ . \ + p_n}.$$

The supposition of

$$p_1 = p_2 = \ . \ . \ . \ = p_n = 1 \tag{17}$$

has, as just noted, not proved satisfactory, extreme values being less trustworthy than moderate ones.

It has been proposed, as an assumption more satisfactory than (17), that each quantity be weighted inversely as its numerical difference from the average, whereby a corrected average will be obtained.[1] Let

$$v_k = x_k - A, \quad (k = 1, 2, \ . \ . \ ., n),$$

then take $\frac{1}{v_k}$ as the weight of x_k, whereby

$$A_i = \frac{\frac{1}{v_1} x_1 + \frac{1}{v_2} x_2 + \ . \ . \ . \ + \frac{1}{v_n} x_n}{\frac{1}{v_1} + \frac{1}{v_2} + \ . \ . \ . \ + \frac{1}{v_n}}.$$

Since the average differs from the centroid on account of $n < \infty$, this corrected average A_i may be treated to a still further correction in the same way. This process, when repeated tends to one of the given measurements as the mean.[2]

What is here approached in a way so awkward as to preclude any but a theoretical interest, can be very simply stated.

Let each quantity have a weight inversely proportional to its difference from the middle value. The centroid of a series of observations thus weighted will coincide with the middle value within the limits of error necessitated by $n < \infty$.

---

[1] ———, *Dissertation sur la recherche du milieu le plus probable*, Annales de math. (Gergonne), 1821 XII 181.

De Morgan, *Theory of prob.*, Encycl. Metropol., II 440, Lond. 1847.

Glaisher, *On the rejection of discordant observations*, Month. Not. Roy. Astr. Soc. Lond., 1873 XXXIII 391.

Stone, *On the rejection of discordant observations*, Month. Not. Roy. Astr. Soc. Lond., 1873 XXXIV 9.

[2] Gergonne, *Note*, Annales de math. (Gergonne), 1821 XII 204.

The case will be represented by a rod without weight having on it a number of particles $x_1, x_2, \ldots, x_n$ with masses inversely proportional to their distance from the point of suspension. The distance of any particle from the point of suspension a is $x-a$; its weight is $\frac{1}{x-a}$; its moment is 1. The point on which the rod will balance is thus determined by the condition that there shall be an equal number of particles on each side.[1] The center of gravity of this system of particles will thus be its middle particle for an uneven number, or a point between the two middle ones for an even number.

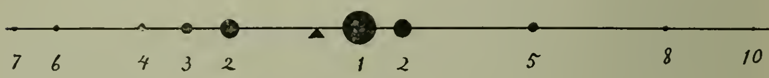

Fig. 4.

For a continuous mass governed by the same law, the centre of gravity x_ξ and the middle value x_m coincide.

From these considerations regarding influence and weight the following conclusions can be drawn:

The median represents the series of quantities in such a way that each quantity has an influence of unity.

The average represents the series of quantities in such a way that each quantity has an influence directly proportional to its numerical value.

The median is equal to the weighted mean where the weights are inversely proportional to the differences of the values from the mean.

The average is equal to the weighted mean where the weights are equal.

Choice of means.

The selection of representative values may take place under three different conditions: I. the results may be so numerous that the empirical frequency curve can be used as the probability curve; II. the results are few but the form of $\phi(x)$ has been determined; III. the results are few and $\phi(x)$ is unknown.

I. *Numerous results.* The fundamental difference between statistics and ordinary measurements lies in the number of measurements executed. Although the passage from one class to another is grad-

[1] Wilson, *Note on a special case of the most probable result of a number of observations*, Month. Not. Roy. Astr. Soc. Lond., 1878 XXXVIII 81.

ual, we can confine ourselves here to the extreme cases. When the results are so numerous that the curve of frequency can be plotted and can be regarded as identical with the curve of probability, every value for x and $\phi(x)$ is assumed as known. The representative values can be calculated from the formulas :
for the centroid

$$x_\xi = \frac{\Sigma\phi(x)x}{\Sigma\phi(x)}$$

for the median

$$\sum_{x_1}^{x_m=M}\phi(x) = \sum_{x_m=M}^{x_n}\phi(x)$$

and for x_h

$$\frac{d\phi(x)}{dx}=0,$$

where $d^2\phi(x)/dx^2$ is $-$.

It is to be noted that there is often more than one value for x_h.

The allowable difference between the actual frequency $\frac{m_k}{m}$ and the probability $\phi(x)dx$ for each value of x can be determined by BERNOULLI's theorem.

When the object of the statistical measurements is merely the determination of the objects measured under a single set of conditions, the result is generally presented in the form of a curve of frequency; thus, all values and their weights being given, there is no need of a selection of any representative value. In scientific work, however, the purpose is generally to determine the change in the results as dependent on a change of conditions; and the observed value is treated as a function of one of the conditions. Given $x=f(z)$ to determine x for each value of z where the measurements for each value of x are numerous. Such an example would be furnished by measuring 10 000 persons each year to determine the law of dependence of height on age. Although the curve of frequency of heights at each year could be made out, still, aside from the impracticability of so much labor in most cases, it would be impossible to give any intelligible expression to the law of relation. Some representative value or values must be picked out for each step of the change. Owing to the fact that x_ξ, x_m and x_h almost never coincide, it is very desirable that all three shall be given. There will thus be three curves,

$$x_\xi=f_1(z),$$
$$x_m=f_2(z),$$
$$x_h=f_3(z).$$

The changes in relative position of these values indicate changes in the character of the quantity measured.[1]

II. *Results not numerous but $\phi(x)$ known.* The law of frequency can be considered to be known: 1. when numerous results have been taken on previous occasions under the same circumstances whereby the law of frequency has been determined with the requisite accuracy; 2. when a knowledge of the circumstances indicates what the law must be.

In this connection it may be well to call attention to the fact that the statement of the law of error as

$$\phi(v) = \frac{h}{\sqrt{\pi}} e^{-h^2 v^2} \tag{11}$$

rests (*a*) on the assumption that the average is the most probable value, (*b*) on the assumption that experience has shown such a law to be true, (*c*) on the fact that it is the limiting form for combinations of symmetrical frequency curves, or (*d*) on the simplicity of treatment thereby rendered possible. I have already pointed out that GAUSS clearly recognized and distinctly stated[2] (*a*) as an assumption, and that long experience has shown (*b*) not to be strictly justifiable. This law of error has done probably better service in astronomy than any other could have done and long familiarity with both assumptions has made them appear almost as axioms. WEINSTEIN,[3] who is careful to call attention to the fact that the assumption (*a*) is not an axiom, is mistaken in asserting that GAUSS regarded it as such. WEINSTEIN is also mistaken in supposing that SCHIAPARELLI[4] attempted to prove analytically that the average is the most probable result. SCHIAPARELLI showed that under certain assumptions the average is the most plausible mean, and stated that it becomes the most probable mean only when (11) is the law of error.

According to FERRERO[5] the utmost defence of the use of the arithmetic mean for all cases is that, when the observations are closely grouped, no mean will differ much from the arithmetic mean.

Apparently still more axiomatic is the law $\phi(-v) = \phi(+v)$. By

---

[1] BOWDITCH, *Growth of children*, XXII Annual Rept. Mass. State Board of Health, 479 (495), Boston 1891.

[2] GAUSS, Theoria motus corp. cœl., II, 3, 177.

[3] WEINSTEIN, Physikalische Maassbestimmungen, I 46, Berlin 1886.

[4] SCHIAPARELLI, *Sur le principe de la moyenne arithmetique*, Astr. Nachr., 1876 LXXXVII 55.

[5] FERRERO, Esposizione del metodo dei mininimi quadrati, Firenze 1876. (I take the statement from a review by PEIRCE, Am. Jour. Math., 1878 I 59.)

most writers it is so regarded. Nevertheless GAUSS makes the statement purely as a hypothesis[1] and LAPLACE gives a special paragraph to the consideration of unsymmetrical facility.[2]

Even if not assumed as an axiom the law is almost universally supposed to have been verified by experience. It is not in place here to consider whether it has been verified for astronomical measurements or not.[3] It has not been verified for psychological measurements. Since the errors of observation in astronomy are in part due to psychological causes, it seems likely that all astronomical records involving an observer would show some assymmetry.

The assumption, without proof, that this law always holds good and that all cases of assymmetry are cases of constant or systematic error, is purely arbitrary.

It is thus evident that many of the cases, supposed to belong in this section, really belong to the following one where $\phi(x)$ is unknown.

According as $\phi(x)$ is (A.) symmetrical or (B.) assymmetrical the treatment of the results and the selection of the representative value will be different.

A. *Symmetrical results.* When the curve of frequency is symmetrical, the ordinate of middle area and the ordinate of the centroid will be the axis of symmetry.[4] Thus $x_m = x_\xi$ and $M = A$ within the allowable limits of error corresponding to the required certainty.

If, according to usual experience, the extreme values occur less frequently than those between the extremes, x_h will in general be the same as x_m and x_ξ, and will be represented by M and A. The curve of probability may, however, have several maxima, none of which may fall at x_m.[5]

In a general fashion the law of frequency for physical, geodetical and astronomical measurements has been found to resemble (11). This law was not, however, originally established on the basis of experience, but was deduced as a necessary result of the arbitrary assumption that A is the most probable value.[6]

Although it has been approximately verified on many occasions, a closer examination shows considerable disagreement in the assymmet-

---

[1] GAUSS, Theoria combinationis observationum, I, 5.

[2] LAPLACE, Théorie analyt. d. prob., 3. éd., 329, Paris 1820.

[3] DE FOREST, *On an unsymmetrical probability curve*, Analyst 1882 IX 135, 142; 1883 X 1, 67 (71).

[4] DE FOREST, *On unsymmetrical adjustments and their limits*, Analyst, 1880 VII 1.

[5] EDGEWORTH, *Observations and statistics*, Trans. Camb. Phil. Soc., XIV 138 (161).

[6] GAUSS, Theoria motus corp. coel., II, 3, 177.

rical position of x_h and in the undue number of extreme values. NEWCOMB even concludes that cases where it is fully valid are exceptional.[1]

In any case of symmetry, whether (11) is valid or not, the median and the average will be theoretically the same.

Since the number of results is small and since according to the principles of probability it is seldom likely that in a small set of measurements the values will be actually symmetrical, the median and the average will frequently differ within the limits consistent with theoretical symmetry. For facility curves of the ordinary exponential form the average is most advantageous[2] as giving a smaller probable and a smaller huge error;[3] for curves very high in the center and widely extended at the extremes the median has the advantage for the same reasons.[4] In neither case is the advantage a great one; in fact, when the ordinary law of probability is assumed, it is practically indifferent which is used.[5]

It is noteworthy that DE FOREST apparently proves that, on the supposition of a symmetrical probability-curve, the influence of the smallness of the number of results renders a symmetrical adjustment of the mean less probable than an unsymmetrical one.[6]

B. *Assymmetrical results.* When the law of frequency is not symmetrical, the vaules x_m and x_ξ can correspond only in those cases where the centre of area falls on the centroid by some peculiar formation of the curve. Such cases, if they ever actually occur, are to be treated as cases of symmetry.

In nearly all psychological and statistical measurements the curve of probability is assymmetrical. If the abscissa of maximum ordinate be determined, the values above it will be found to be much more frequent than those below it. That is, $x_m > x_h$. For all assymmetrical curves of this general form, as the assymmetry increases, x_ξ departs more rapidly than x_m from the main mass of results,[7] and consequently does not represent them so well.

The general expression for the usual cases of assymmetry has been

---

[1] NEWCOMB, *A generalized theory of the combination of observations*, Am Jour. Math., 1886 VIII 343.

[2] LAPLACE, Théorie analyt. des prob., 2 Suppl., § 2.

[3] MERRIMAN, Method of Least Squares, New York 1894.

[4] EDGEWORTH, *Observations and statistics*, Trans. Camb. Phil. Soc., XIV 138 (167).

[5] EDGEWORTH, *Choice of means*, Phil. Mag., 1887 (5) XXIV 268 (270).

[6] DE FOREST, *On an unsymmetrical probability curve*, Analyst, 1883 X 67 (74).

[7] DE FOREST, *On an unsymmetrical probability curve*, Analyst, 1883 X 67.

deduced by De Forest.[1] It includes constants determined from the squares and cubes of the errors. Any method, however, that introduces more calculation than the usual average will be at even greater disadvantage than that value.

If Fechner's law of the estimate of differences of sensation could be relied upon in all cases, the best representative value would unquestionably be the geometric mean. If the geometric mean g be assumed as the most probable value, it is easily shown that

$$\phi(x) = \mathrm{B}e^{a\left(\log \frac{x}{g}\right)^2}$$

which with the usual assumptions becomes[2]

$$\phi(x) = \frac{h}{\sqrt{\pi}}\, e^{-h^2\left(\log \frac{x}{g}\right)^2}. \qquad (19)$$

The general result of experience, however, goes to show that $\phi(v)$ is of a form intermediate between (11) and (19).

For cases of assymmetry the most natural representative value to take would be x_h. As this is not determinable from few results, some other value must be used. Any other value would be justifiable only from a consideration of the purpose for which it is wanted. There is no reason, as far as I can see, for taking any one of the values around the maximum rather than any other one except in so far as it comes nearer the maximum. If we are to take x_m or x_ξ and not x_h, the middle value x_m would be the better on account of its nearness to x_h.

III. *Few results and unknown* $\phi(x)$. Since it is impossible from the few results given to make any deductions concerning $\phi(x)$, it is evident that other things being equal, that value will be preferable for which the fewest assumptions need to be made.

Let it be assumed that the curve of frequency is symmetrical. Then $x_\xi = x_m$. One value is as good as the other, for both should be the same. The value of maximum probability x_h cannot be directly calculated. For any assumed form of $\phi(x)$, a comparison of the mean variation, the mean-square-error and the probable error will

---

[1] De Forest, *On an unsymmetrical probability curve,* Analyst, 1882 IX 135, 161; 1883 X 1, 67.

[2] McAlister, *On the law of the geometric mean,* Quart. Jour. Pure a. Appl. Math., 1880 XVII 175.

McAlister, *The law of the geometric mean,* Proc. Roy. Soc. Lond., 1879 XXIX 367.

show whether they stand in the proportions required by the assumption for $\phi(x)$.

Since $\phi(x)$ is in general unknown, or only roughly suspected, it becomes desirable to either occasionally or constantly compare x_ξ and x_m in order to judge of the symmetry or assymmetry of $\phi(x)$.

For a symmetrical curve $x_\xi = x_m$. If for a supposed symmetrical curve over $x_\xi = x'_m$ the values r', s' and μ' have meanings as defined on p. ? and if r, s and μ be the corresponding values for x_m, it can be expected, according to Bayes's theorem, with a probability of

$$\Phi(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2}\, dt$$

that

$$\frac{r'}{\mu'} = \frac{r}{\mu} \pm \gamma \sqrt{\frac{2r^2}{\mu^3}}$$

or since

$$\frac{r}{\mu} = \frac{1}{2}$$

then

$$\frac{r'}{\mu'} = \tfrac{1}{2} \pm \frac{\gamma}{2} \sqrt{\frac{2}{\mu}}.$$

Instead of x_ξ and x_m we have $M = x_m \pm \eta$ and $A = x_\xi \pm \epsilon$ where η and ϵ are determined by (12) and (13). It would not be difficult to calculate on general principles the limits of difference between M and A within which we could, with a given degree of probability, suppose that the law is symmetrical. For the present I will assume that the desired degree of probability is $50\% = \frac{1}{2}$ and that the question to be decided is the symmetry or assymmetry of a curve whose equation is given in the case of symmetry by (11).

If the curve be symmetrical around the average, the probable error R of the variations

$$x_k - A, \qquad (k = 1, 2, \ldots, n)$$

will be the limit of variation for A. If the curve be symmetrical around the median, M can be considered as representing the average of such a curve; the probable error R′ calculated from

$$x_k - M, \qquad (k = 1, 2, \ldots n)$$

will give the limit of variation for M. Let the difference between the average and the median be denoted by

$$\delta = A - M.$$

As long as the limits of R and R′ overlap, that is, $\frac{R+R'}{2} < A - M$, the presence of assymmetry cannot be asserted. But when $\frac{R+R'}{2} > A - M$, it can be said with a probability of 50% that the curve is assymmetrical.

If M and A indicate symmetry in a large number of cases, either of these values can be used in similar cases for the reason that both are practically the same.

If they indicate assymmetry, there are the same reasons for prefering M to A as in the case of known $\phi(x)$ considered above.

To Profs. Gibbs, Newton and Elkin of Yale and Prof. Merriman of Lehigh, I am under very great obligations for discussion, criticism and correction. Of course, they are in no wise responsible for my deductions or conclusions, from which each one dissents at some point. Nevertheless, any value this article may have is due to their patient labor with one who is not a mathematician but who is obliged to use mathematical means to solve practical problems.

RESEARCHES ON THE MENTAL AND PHYSICAL DEVELOPMENT OF SCHOOL-CHILDREN,

BY

J. Allen Gilbert, Ph.D.

Infant development during the ages from 3 to 6 has been treated rather extensively, but so far as systematic scientific work is concerned, very little has been done in the study of the child during the years spent in school except in the line of bodily growth in weight, height, chest-capacity and the like. In measuring mental processes almost nothing has been done. The work of Bolton[1] on the growth of memory in school-children has brought out a valuable subject but would have been of far greater value had its statistics, which were obtained from careful work, been presented in more intelligible form. The subject of voluntary motor ability has been treated by Bryan.[2] Hearing has been investigated by Chrisman.[3]

Last year I carried through a series of tests on school-children to determine their sensitiveness to differences in pitch,[4] and through this and the problems it suggested I was led to continue my investigations on a much larger scale in order to aid in that analysis of mental phenomena which is so necessary to an understanding of child-psychology. The present investigation was undertaken with the determination to carry through a regular set or series of accurate mental and physical tests upon school-children from 6 to 17 years of age.

The work has occupied most of my time during the academical year 1893–1894, the larger part of the fall term being spent in the invention and construction of apparatus used in taking the tests. In the furtherance of my work I am specially indebted to the following persons: to Dr. E. W. Scripture, who has charge of the Laboratory and was always ready to assist any endeavors at honest work; for

[1] Bolton, *Growth of memory in school-children*, Am. Jour. Psych., 1893 IV 919.

[2] Bryan, *On the development of voluntary motor ability*, Am. Jour. Psych., 1893 V 123.

[3] Chrisman, *The hearing of children*, Ped. Sem., 1893 II 397.

[4] Gilbert, *Experiments on the musical sensitiveness of school-children*, Stud. Yale Psych. Lab., 1893 I 80.

his assistance I am truly grateful; through his suggestions I was led to take up this investigation. To Mr. V. G. Curtis, superintendent of schools of New Haven, for the permission to enter the schools with my tests and for his kindly interest. I wish also to express my gratitude for the kindness and accommodation offered by Miss Webster, Messrs. Camp, Hurd and Thomas, principals of Welch, Dwight, Winchester and High Schools respectively; to Miss Treat, assistant at Welch for the time spent on tests (1) to (3), together with all the teachers at the respective schools who were helpful during the months occupied in taking the tests; to Dr. C. B. Bliss, fellow and assistant at the Psychological Laboratory, for much of his time and many suggestions in the preparation of apparatus and prosecution of my work throughout; to Mr. Hogan, the laboratory mechanic, for his assistance in the construction of apparatus. To Professor Williams I am indebted for the idea embodied in the suggestion-test. Valuable suggestions were also received from Professor Ladd, who has throughout shown a kindly interest in my work.

Methods and apparatus.

Each child was tested in the following respects: muscle-sense, sensitiveness to color-differences, force of suggestion, voluntary motor ability, fatigue, weight, height, lung-capacity, reaction-time, discrimination-time and time-memory. Ten sets of special apparatus were constructed for each of the first three tests.

Test (1): *Muscle-sense.*

In this test each set of the apparatus consisted of ten weights varying from 82^g to 100^g in steps of two grams each. The weights were brought to the exact weight within 50^{mg}. In order to get at mere muscle-sense or sensitiveness to weight, which is the aim of this test, heat, cold and roughness, which would distract the attention from this one sensation of weight in consciousness, must necessarily be avoided. To avoid heat and cold, cartridge shells were used, giving a paper surface, which is a non-conductor. Uncapped shells were cut in two, giving a cylinder 2.3^{cm} in diameter and 3.8^{cm} long. These were filled with lead disks and brought to within 100^{mg} of the weight desired. In order to avoid sensations of roughness, they were painted with asphalt, which leaves a hard glazed surface, raising the weight to within 50^{mg}. A cylinder of the size named and filled as described makes a weight comparatively heavy and still of

sufficiently small size to be grasped easily endwise between the thumb and finger by a six-year-old child. In cutting the lead disks with which the shells were filled, the punch was so constructed as to press the disk into convex-concave shape before cutting it from the sheet of lead. These were placed with concavity to concavity and on being pounded down of course flattened out and consequently became of diameter large enough to press firmly against the side of the shell, thus avoiding any jostling sideways within, while they were also bound tightly one by the other, thus avoiding any jostling endwise. To get them of different weights sufficient cartridge-wads were used in place of lead disks. Each weight was marked by a secret sign to indicate its weight. For each set a box was then made of appropriate size in order to avoid any mixing of the sets in taking the tests.

In taking the tests the lightest one, marked by a white speck on the end and weighing 82^g was used as the standard. The child received the box of weights and was told to sort out all those which seemed to him to be of exactly the same weight as the one with the white speck on the end, lifting them endwise between thumb and finger. To avoid the effects of fatigue on the sensitiveness to weight, each child was given only two trials on each block, lifting it and the standard alternately. The successive steps between the weights being two grams, the number of blocks selected as being of the same weight when multiplied by 2 would indicate in grams the threshold for discrimination to weight for that child.

Test (2): *Sensitiveness to color-differences.*

Just as in the preceding test we were in search of the threshold for discrimination to weight or least perceptible difference in weight, so, in this test it was the aim to find the threshold for discrimination to color or the least perceptible difference in shade of one color. This test consisted of a series of ten shades of red so closely graded that no two successive colors or shades could be distinguished except by an experienced eye. Gray would have been preferable and in fact was tried previously to red, but owing to the fact that all goods are bleached with sulphur, no matter how well scoured before dyeing, traces of red could be found running through the gray.

Ten pieces of woolen cloth of fine texture were first dyed a suitable red by a practical dyer under my supervision. After the ten pieces were removed from this coloring solution, which left them all exactly

of the same color, a very small portion of dye was added to the boiling vat, thus making the fluid slightly darker. One of the pieces was again boiled in this, making it, when removed, very slightly darker than before. To this last solution was again added a small portion of dye in which a third piece was boiled and so on, adding for each successive piece of cloth an equal portion of dye and giving thus a series of shades each differing from the others in a very slight degree. Each of this series of ten was then fixed firmly in a ring so as to exhibit the color and yet protect it from being handled. This ring was a hollow cylinder with a narrow shoulder on one end, the edge of which was beveled so as to avoid any shadows being thrown on the color. The colored cloth was stretched firmly across the end of a circular block which was driven into the ring, holding the color firmly against the shoulder. The circle in which the color was exposed was 3^{cm} in diameter. The castings were then painted a dull black so as to avoid any reflection of light which might affect the color. Each block was then marked with a secret mark according to the place it held in the series. To avoid getting them mixed in taking the tests, a small box was made for each set of ten.

In taking the tests the block containing the lightest shade, painted white on the bottom to distinguish it, was used as the standard with which to compare the rest. The child was given a box containing one set and told to pick out all those shades of red which were exactly like the one painted white on the bottom. The number of those selected as being alike, including the standard, was then recorded. The number of colors picked out would indicate 'the threshold for discrimination to color for that child and by averaging the individual results the discrimination for the respective ages was obtained.

Test (3): *Force of suggestion.*

The aim in this test was to measure the effect of our ideas of a thing formed by the sense of sight upon those formed by the muscle-sense, in particular to get the effect of bulk in a thing upon our judgment of what it should weigh. As a first attempt, a set of ten blocks was made, each weighing 55^{g} with an accuracy of 50^{mg}. All were 2.8^{cm} thick but varied in diameter from 2.2^{cm} to 8.2^{cm} in a geometrical series, in order to make the sensations increase in an arithmetical series according to Weber's law. On taking a number of tests with this set, asking the subject of the

experiment to arrange them in order according to their respective weights, the decision was universal that the smallest one seemed heaviest and the largest one lightest, the others ranging between those two extremes with weights inversely proportional to their size. Now, in order to measure the amount of the suggestion offered by the bulk of the blocks, all being of the same weight in reality, a series of fourteen blocks was made, each being 2.8^{cm} long and 3.5^{cm} in diameter but all of different weights, ranging from 15^{g} to 18^{g} in weight. In order to get the different weights in the same sized blocks, holes were bored out of the center of a size directly proportional to the weight desired and filled with lead. This could then be easily bored out till the exact weight for each was reached. The lead was then concealed by a cartridge-wad, which also served the purpose of a center to the block on each end by which to grasp it when lifting. The position of each block in the series of fourteen was then marked plainly on one end.

In taking the tests for measuring the amount of suggestion offered by the difference in bulk, the very large weight and the very small weight were given to the child in connection with the series of fourteen blocks, the weight of the large and small standards being unknown to the child. He was asked to pick out of the fourteen blocks the one that seemed to him to be of the same weight as the small standard and also the one of the same weight as the large standard, lifting them endwise between the thumb and finger. The small block was first lifted; then the end-one of the 14 blocks was lifted. If it was apparently lighter, the second block of the 14 was tried. If this was also lighter, the third was tried. This was continued till that block of the 14 was reached which was apparently equal to the small block. Its number was noted. Then the large block was compared successively with those of the 14 till an apparently equal one was reached. The weight of the small one was the same as that of the large one. The amount of suggestion offered by the difference in bulk, could then be measured in grams by taking the difference in weight between the two blocks chosen as being the same weight as the large and small standards respectively. One heavier than 55^{g} was always chosen for the small one and one lighter than 55^{g} was always chosen for the large one, as will be explained in the discussion of results.

The first three tests just described were taken on desks adapted to the size of the child, thus subjecting all to the same influences in lifting. The desk was of such height as to throw the fore-arm paral-

lel with the floor. These first three tests were given to the child in the order in which I have recorded them, thus throwing the color-test in between the two weight-tests, giving no chance for the fatigue of test (1), should there be any, to be carried over into test (3).

TEST (4): *Letter-memory.*

This was omitted because of the impossibility of accurate work with letters.

TEST (6): *Weight.*

The weight of the children was taken on a balance-scale weighing with an accuracy of one-quarter of a pound. Ordinary in-door clothing was worn.

TEST (7): *Height.*

The SEAVER measuring-rod, marked off in both inches and centimeters, was used. It is composed of a straight stick with a sliding arm projecting at right angles. Placing the stick perpendicular to the floor by sighting it parallel with a door-frame, the sliding arm was made to touch the head of the child and then read off in tenths of a centimeter, as marked on the stick. The height was taken with shoes.

TEST (8): *Lung-capacity.*

The Standard wet spirometer was used. This consists of a cylindrical vessel nearly filled with water, through the center of which a tin tube projects, connected at the lower end with a rubber tube through which the experimentee exhales the air. Over this tin tube, projecting in the center of the vessel of water, a tin cylinder, closed at one end, is inverted and allowed to sink in the water by opening a stop-cock at the bottom, letting out the air confined in the vessel by the water. An index finger pointing to a scale on a support on one side, marked off in cubic inches, is fastened to the movable cylinder. As air is blown into the tube the hollow cylinder rises, marking off on the scale the number of cubic inches blown into it. The weight of the tin cylinder is balanced by weights hanging on pulleys above, so that very little pressure is required to raise the cylinder by blowing.

The child was told to inhale into its lungs all the air they possibly could contain and then to exhale it through the tube into the spirometer, emptying his lungs as completely as possible. The cubic inches were reduced to cubic centimeters for the final averages of each age.

The reaction-board.

Since mental tests on children have never been taken to any great extent, there was consequently no suitable apparatus at hand for such tests. The carrying out of my experiments necessitated the construction of what may be called the reaction-board, arranged for taking tests (5), (9), (10) and (11). This was constructed on an oak board 33 centimeters square, fig. 5. The main parts are the electro-

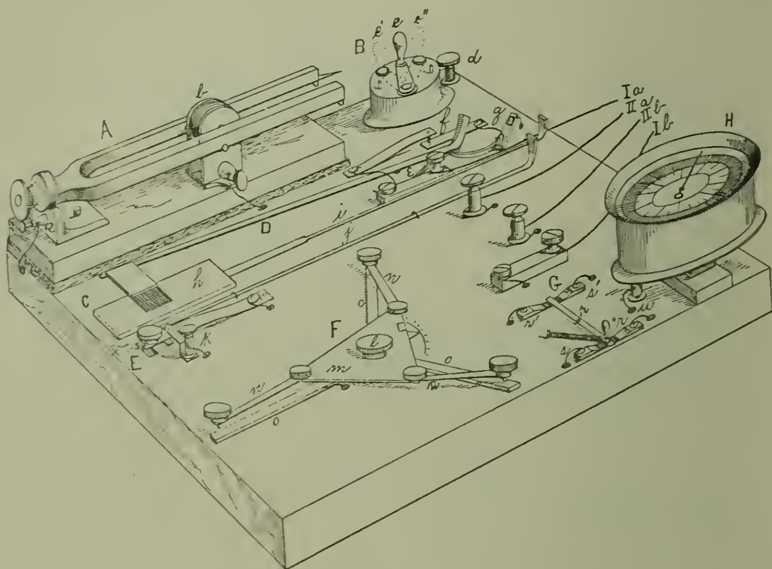

Fig. 5.

magnetic tuning-fork *A* vibrating one hundred times per second, the double-post switch *B*, the stimulating apparatus *C*, the reacting-key *E*, the tapping-apparatus *F*, the commutator *G* and the Ewald chronoscope *H*. The board is raised from the table by four short legs so as to permit insulated wires to pass beneath connecting the different parts of the apparatus through holes piercing the board. One leg *d* is an adjustable screw by which the board can be fitted to any surface upon which it may be placed. Two separate Grove batteries had to be used, one connected with the wires I*a* I*b* and the other with II*a* II*b*. Current II*a* II*b* simply passes through the tuning-fork *A*, when the bar *p* of the commutator is left in the position in which the spring naturally holds it, viz: connecting the two posts *s* and *s′*. Thus, the tuning-fork is kept in constant motion.

The chronoscope is composed of an electro-magnet with the armature connected with a small lever, one end of which rests against a toothed wheel connected with the finger on the dial visible at *H*. Every time a current is made to pass through the electro-magnet it draws the lever, thus moving the wheel and the finger on the dial one mark. The circle on the dial of the chronoscope is divided into one hundred parts and thus, if the chronoscope is thrown into a current connecting it with the tuning fork *A*, which vibrates one hundred times a second, the finger on the dial makes one complete revolution in one second. Every time the fork vibrates the current is made at *b* by means of an adjustable wire invisible in the figure. Every time this current is made at *b* the chronoscope moves one mark and thus records the number of vibrations made by the fork, or, in other words, measures in hundredths of a second the length of time a current is allowed to pass through it and the fork. In order to test the chronoscope it was thrown into circuit with a time-marker on a smoked drum according to a method described by Bliss.[1] To verify results this test was made both before and after the taking of data. The chronoscope was found accurrate to the error of scale. Owing to the fact that the contact between *g* and the stimulating rod *D* is a "make" contact, an error of 0.005 of a second was introduced, but in as much as the chronoscope only records in hundredths of a second, this error would not influence the results.

The apparatus *F* which, for my present use, I have called the tapping-apparatus, was at first intended to be used as a habit-key. By unscrewing *l* and the screws binding the three arms *n* to the central equilateral triangular plate *m*, the plate *m* could be turned to the right thus drawing in toward the center the buttons on the ends of the bars *n*. The different radii of the circle were measured off on the scale on the face of the board as the index moved to the right or left, according as the imaginary circle, passing through the three buttons on the end of the arms *n*, was desired smaller or larger. To obtain an expression for habit, the child could be told to tap on the three buttons going in a circle to the right as fast as possible, for five or ten seconds, the number of taps being recorded by the chronoscope. After resting he could then tap the same length of time in the same circle. The increase in number of taps, or percentage of gain, in the second trial over the first, would express the rapidity of forming the muscular habit of moving the hands in a set way. This

---

[1] Bliss, *Researches on reaction-time and attention*, Stud. Yale Psych. Lab., 1893 I 5, fig. 3.

test was given up because it seemed impossible to get the children to keep on going in a circle should they miss one of the buttons; instead, they would almost invariably stop and try to correct their misdirected aim. Although this, however, would make an interesting test on adults, it lay outside of my problem.

Test (5): *Voluntary motor ability and fatigue.*

By throwing the arm c of the switch B to the post covered by it when in the position c' the current Ia Ib is made to pass through the wire Ia across to the reaction-key E, to the binding post g, then through c in position c' to the tapping apparatus F; thence through the keys n and bars o to the chronoscope H and finally back to the binding post Ib.

In measuring voluntary motor ability, the child was asked to tap as rapidly as he could, until told to stop, on the button at the end of the front key n. Closing the key f closed the current at every point except at the platinum contact between the key n and the bar o. Thus, every time the child tapped on the button of key n the contact was made with the bar o allowing the current to pass through the electro-magnet of the chronoscope, moving the finger on the dial one mark. An upright circular screen was fastened to the edge of the dial in order to hide from the child the record made. The child tapped for forty-five seconds. Shortly after he had started tapping, the circuit was closed by pressing down the key f in unison with one of the strokes of a metronome which was adjusted to beat seconds. As soon as the key f was pressed down by myself, the chronoscope commenced recording on its face each tap made by the child. At the end of five seconds I broke the current at f thus cutting off from the chronoscope any means of recording the taps of the child until the current was again made. The child continued tapping simply to produce fatigue. At the end of forty seconds I again made the circuit at f and took a record of the taps made in the last five of the 45 seconds in the same way that the first five were taken. To measure off exactly 5 seconds I counted 0, 1, 2, 3, 4, 5 in unison with the beat of the metronome, pressing down the key f at 0 and releasing it at 5, thus giving an interval of 5 seconds. By this means the control of the 5 seconds was in my own hands; this avoided all such errors as are sure to creep in when the child is simply told to start and stop tapping at word of command, as was done by Bryan in his researches on voluntary motor

ability.[1] In such a case it is very difficult to tell just where the counting of taps should cease, for almost invariably a child adds a stroke or two after being told to stop. By the method used here, however, the tapping is limited to exactly 5 seconds. By throwing the chronoscope into a current with the tuning-fork, vibrating one hundred times per second, I contrived a plan to measure my own accuracy in making and breaking the current at *f* in unison with the metronome, leaving exactly 5 seconds intervening. In 5 seconds the hand on the dial of the chronoscope should revolve 5 times. After taking and averaging fifty trials at making a 5-seconds-interval, it was found that my average variation for a single occasion in making exactly the right length of time was only 0.02 sec. Such an error would be wholly negligible since the highest rate of tapping obtained was 47 taps in 5 seconds. The number of taps made in the first 5 seconds can be taken to represent the voluntary motor ability of the child. It is impossible to tap as rapidly after 45 seconds as at first; by calculating the difference between the two rates of tapping and then dividing this difference by the number of taps made the first 5 seconds, an expression for fatigue was obtained as a per cent. If r be the number of taps for the first 5 seconds and s the number for the last 5 seconds, the degree of fatigue for 40 seconds of tapping can be expressed by

$$g = \frac{r - s}{r}.$$

To have expressed the fatigue merely by the difference between the two rates of tapping would not have expressed the truth ; for instance, one child who tapped 19 and 15 for the respective periods of 5 seconds, lost a great deal more by fatigue than another, who tapped 38 and 34 respectively; each lost 4 taps but the first lost 21 per cent., the second only 11 per cent.

Since fatigue was the principal problem I had in view in this test, the elbow was held free from the table so as to bring on fatigue the more rapidly. This also would be the most rapid way of tapping, for it consists largely of a movement of the wrist which is one of the most rapid joints for tapping.[2]

Test (10): *Reaction-time.*

By throwing the arm c of switch B, fig. 5, to the post covered when in position c'' current Ia Ib passes, when all points are connec-

---

[1] Bryan, *Voluntary motor ability*, Am. Jour. Psych., 1893 V 14.

[2] Bryan, *Voluntary motor ability*, Am. Jour. Psych., 1893 V 171.

ted, from the battery through wire I*a*, through the spring *e* to the end of the rod *D*, thence through the curved spring to *g*, thence to and through the arc *c* of switch *B*; thence to key *e* through binding post *k*; thence to *r* of the commutator through rod *p* when thrown across from *r* to *r'*; thence through the chronoscope, back again to the battery. Current II*a* II*b* is connected with the following in succession: II*a* from battery, binding post *a* on tuning-fork *A*; binding post on the coil of the fork; *s* and *s'* of commutator *G*, finally, going back to the battery through wire II*b*. The coil-spring on rod *p* keeps it continually in connection with the poles *s* and *s'*, thus keeping current II*a* II*b* always closed and the tuning-fork in constant motion, ready for use. Between *r* and *s*, *r'* and *s'* are two small pieces of hard rubber to insulate and carry the arm easily from *ss'* to *rr'*. With the arm *D* thrown back by the spring *e* against the box *h* the contact is merely broken between the end of the arm *D* and the curved spring fastened to the brass block *g*. The child is told to press down the key *E* as soon as he sees a movement of the disk fastened to the end *D*. By throwing the rod *p* of the commutator *G* upon *r* and *r'*, the current I*a* I*b* is closed at every point except where the contact is broken between the arm *D* and the curved spring on *g*. By throwing the arm *D* into the position seen in the figure this contact was made, completing the circuit and at the same time giving the stimulus for reaction. Throwing the commutator into position *rr'*, before throwing the stimulating rod *D*, served as the warning to the child that the stimulus would come, as well as changing the current used, from II*a* II*b*, which passes through the tuning-fork alone, to the current I*a* I*b* which passes through tuning-fork *A*, stimulating rod *D*, reacting key *E* and chronoscope *H*. As soon as contact was made at *g*, by throwing the stimulating rod *D*, the current, being made and passing through tuning-fork *A*, which is vibrating 100 times per second, started the chronoscope going at the rate of 100 marks or one revolution per second. As soon as the break-circuit key *E* is pressed, the current is immediately broken; thus stopping the chronoscope. This records the number of hundredths of a second which elapsed between the movement of rod *D* by myself and the pressing of the key *E* by the child. After the reaction, the key *E* is grasped and held down by a small spring *x* until the record can be read from the chronoscope and recorded. Spring *e* also throws the rod *D* back into its original position, breaking the contact at the spring on *g* as before. After a record is made on the card of the number of hundredths of a second it took the child

to react, the spring x is loosened by pushing it off the key E by the rod j and the finger of the chronoscope is moved back to 0 ready for the next test. Each child was given ten trials, the median value of which was taken for his reaction-time. In order to prevent the child anticipating the moment when the stimulating rod D is to be thrown by seeing the hand move, a screen extended back over my hand and the apparatus about the part e, f and g. This screen has been removed in the figure so as to exhibit the keys e, f and g. The face of the chronoscope H was also hidden from the child by a small semi-circular screen fastened to its edge. The Ewald chronoscope,[1] as seen in the figure, is taken from its usual stand and fastened to the board so as to place it near the commutator, thus allowing it and the commutator to be manipulated conveniently with the left hand.

Test (9) : *Reaction with discrimination and choice.*

The apparatus for this test was the same as in simple reaction with the addition of the color apparatus h with rod i. Fastened to the rod i and concealed in the box h is a slide nearly as wide as the box but only two-thirds as long. Upon this slide are glued two pieces of colored paper, red and blue, each taking up one-half of the slide. By pushing in the rod i till it strikes at the end of the box at C, blue is exposed at the opening in the top, while red is concealed under the top of the box between the opening and the end marked C. By pulling out the rod i till it strikes the other end of the box h, red is exposed. The exposure is made when the stimulating rod D is thrown into the position seen in the figure. The child was told to react on the key E if the color, when exposed, was blue, and not to react if it was red ; this compelled him to wait and discriminate between red and blue and also to make the choice whether to react or not. First, the bar p of the commutator G is thrown from current ss' to rr', serving also as a warning to the child that the stimulating rod D will soon be thrown aside. The rod D is then thrown, whereby the chronoscope is started. Pressure on the key E immediately stops the chronoscope while the spring x also holds down the key E until the number of hundredths of a second, counted off by the chronoscope, can be recorded on the record-card. After moving the finger of the chronoscope back to 0 by a wheel invisible in the figure and releasing the spring x by pushing rod j, the apparatus is ready for the second test, the commutator p having been brought back to ss'

[1] Dumreicher, *Zur Messung der Reactionszeit*, Inaug. Diss. Strassburg 1889.

again by the spring attached to it. That part of the rod *i* which is not concealed by the screen over the hand and keys *e*, *f* and *g*, is concealed by a wooden covering so as to prevent the child from knowing by the position of the rod what color will appear. For reasons mentioned below this test preceded that for reaction-time.

Test (11) : *Time-memory.*

The same apparatus on the board was used here as in the preceding test, except that instead of using the stimulating rod *D* to start the chronoscope, the key *f* was used. The chronoscope, when in motion at the rate of one hundred marks a second, makes the same tone as the fork only somewhat louder and of different timbre. The tuning-fork is mounted on hair-felt so as to muffle its sound, leaving the sound of the chronoscope very easily heard. The child was told to listen how long I caused the chronoscope to sound and then after I started the sound the second time he was to stop it by pressing down the key *E* when he thought it had gone just as long as I allowed it to go the first time. After throwing the commutator from *ss'* to *rr'* to change currents and also to warn the child that the sound would soon begin, the current was made by pressing down the key *f*, which was held down till the finger on the dial of the chronoscope *H* had made two complete revolutions, whereupon it was released ; thus the sound continued for two seconds. The shade concealing the dial kept the record from being seen. Almost immediately after the current was broken, it was again made at *f*. This again started the sound which continued until the child stopped it by pressing down the key *E*. The key was held down as before by the self-catching spring *x* until the record on the chronoscope could be read and recorded. The figures entered on the card represented the error, in hundredths of a second, made by the child in trying to make the second sound just as long as the first. As it was impossible for me to make the standard exactly two seconds, this error, never more than 0.05 of a second, was added to or subtracted from the error of the child according to its direction. Each child was given ten trials from which the median value was calculated for his general result.

General methods.

Tests on muscle-sense, color-sensitiveness and force of suggestion were not taken simultaneously with the other tests. In taking the tests on voluntary motor ability, fatigue, weight, height, lung-capacity, discrimination-time, reaction-time and time-memory the

following order was adopted. Three children were taken from the school-room at one time into a secluded room away from interruption, noise, etc. While one was taking the tests the other two could be watching and thus, when their turns came, they understood what was expected of them, and took but little time for explanations. Furthermore, when their turns came the novelty of the board and tests was worn away, so that their whole attention could be devoted to doing the things required. The child was first weighed, his height was taken next and then his lung-capacity was measured. After this he was subjected to the tests of the reaction-board, the first being discrimination-time for red and blue. This test was given precedence in time to simple reaction because after getting accustomed to reacting every time the stimulating rod was moved, as is the case in the simple reaction, it is very much more difficult to refrain from pressing reflexly or automatically when the red is exposed in the discrimination-test than if the latter is placed first. In this test the number of errors made by pressing the key E when red was exposed, instead of not pressing at all, was kept account of. The red and blue were exchanged irregularly so that the child could get no idea of what color to expect. So as to put him more thoroughly on his guard, several reds were generally exposed at the start. Ten trials were given in this and in the two succeeding tests, the result of each trial being recorded in its proper place on the card as soon as taken. Reaction-time was taken next, and then the test on time-memory. Finally, after all other tests had been taken (which required from seven to ten minutes) the tests on voluntary motor-ability and fatigue were taken. During the 35 seconds interval between the first and last period of five seconds of the forty-five seconds of tapping, the next child's weight and height were taken. When the child at the board had completed the fatigue-test he was sent back to the room and immediately another came to fill his place, thus keeping three out at one time. Just before taking the reaction-board-tests on the child the day of the month and the hour of the day were noted after "date" on his card so as to enable me to calculate at some future time the effects of weather and also of fatigue produced by the day's work upon those tested late in the day, compared with those taken early in the morning. The teachers attended to having the remainder of the data filled out at the head of the card, in regard to the child, namely, age at last birthday, birth-place, birth-place of father, birth-place of mother, father's occupation. White cards were used for girls and colored cards for boys.

Results.

About one hundred children of each age were taken, the small variation from this number being shown in column N of the tables I to XI. The method explained above necessitated tests one to three inclusive being taken at a different time from the remainder. Some of the children, having taken one portion of the tests, were absent at the time the other tests were taken, thus causing the number to fall slightly below one hundred in some instances. The results for each age were averaged into one final result by taking the median value[1] according to the formula $\frac{n+1}{2}$. The justifiableness of this method will be shown later both by data and curves.

Fechner has proposed the use of the median or central value, whose position in the series of separate results arranged according to size is given by $\frac{n+1}{2}$. That is, if all the results are to be arranged in the order of their size, the median will be just in the middle. Since with finite units of measurements there will be a number of results having the same value around the middle, the value will be determined by interpolation. The importance of the use of the median lies in the fact that it involves no assumption in regard to the distribution of the separate deviations.

In order to get an expression for the homogeneity of my results the mean variation was calculated for each age. This same calculation was also made for boys and girls separately in all the tests except the first three, viz: muscle-sense, color-sensitiveness and force of suggestion. After the record-cards were completely filled out with the data desired, they were returned to the teachers of the respective rooms, who were asked to mark each name by a figure 1, 2 or 3 according to what she judged the child's general mental ability to be, marking the bright ones 1, those of average ability 2 and the dull ones with a figure 3. The aim of this was to get the relation between the general mental ability and the respective tests. The tests were taken from January 17th to April 1st, 1894.

Test (1): *Muscle-sense.*

The results for this test are shown in table I and the accompanying charts I and II, giving in graphic form the results as given in columns D, B, G and MV of the table. Sensitiveness to weight

---

[1] Scripture, *On the adjustment of simple psychological measurements*, Psych. Rev., 1894 I 281.

was not so delicate as was supposed when the apparatus was made, and consequently a series of ten weights varying two grams each, from 82^g to 100^g, was insufficient to include all of the younger children. In all ages, except fourteen and fifteen, one or more children were found whose sensitiveness to weight-differences did not fall within the scope of my apparatus, viz. 18^g variance from the standard weighing 82^g. In calculating column *D* of table I, all data including those in which all ten were said to be alike in weight were used. In order to correct this error of apparatus, the per cent. of data in which all ten weights were reported as being alike was calculated for each age. To obtain a correct estimate of the discrimination to weight for each age, the columns *D* and *P* must be considered together; that is, the results for this sense give the wrong impression, unless both of the columns be considered in conjunction. The figures at the left of the chart indicate the threshold for discrimination to weight in grams. The figures found in columns *P*, *PB* and *PG* of table I represent in per cent. the number of children who picked out all ten weights as being exactly alike, distinguishing no difference whatever in their respective weights. It will be remembered that sensitiveness varies inversely as the size of the least perceptible difference. For example, the least perceptible differences for the ages 6 and 7 are 14.8^g and 13.6^g respectively; the threshold for 6 yrs. bears to that for 7 yrs. the relation $\frac{14.8}{13.6}$, but the sensitiveness for 6 yrs. is greater than that for 7 yrs., the relation being $\frac{13.6}{14.8}$. In general it is convenient to indicate the sensitiveness by the reciprocal of the least perceptible difference, thus, $\frac{1}{14.8}$, $\frac{1}{13.6}$, etc. On the chart, the higher the line the greater the least perceptible difference but the smaller the sensitiveness. Ages are marked along the axis of abscissas at the bottom of the chart.

The results show a gradual increase in ability to discriminate, from the ages of 6 to 13. At 6, the worst year of any for discrimination, the least perceptible difference was 14.8^g, with 38% of non-discriminations; at 13 years only 5.4^g with 2% of non-discriminations. After 13 there was a gradual falling off of 6.8^g, none failing to discriminate, and then another gain till at 17 it was 5.8^g with 1% of non-discriminations. Boys and girls, considered together, gradually increase in ability, but when they are considered separately, marked differences of sex appear. At 6 there is the large difference of 3.8^g in discriminative ability in favor of the boys. At 7 they have the same ability. From this on, they gain with equal pace to the year 13 with the exception of the abrupt falling off for boys at 11. From 13 to 17

the difference in ability again becomes manifest in favor of boys. In general it may be said that the superiority of boys in sensitiveness to differences in weight increases with age, irregularities being noticeable, however, from 6 to 7 and from 12 to 14.

It is interesting in this connection to notice the relation between the general curve, chart I, and the curve of mean variation for the same test, chart II. There seems to be a general agreement throughout between the main curve for discrimination and the curve for

TABLE I.

Muscle-sense.

Age.	*D*	*P*	*MV*	*B*	*PB*	*G*	*PG*	*N*	*NB*	*NG*
6	14.8	38	5.2	13.0	26	16.8	49	87	42	45
7	13.6	36	4.4	13.2	36	13.2	40	92	50	42
8	11.4	30	4.6	12.2	35	11.0	28	92	46	46
9	10.0	20	4.4	10.2	23	10.0	17	95	48	47
10	8.8	12	4.4	8.6	12	9.2	12	91	49	42
11	8.6	6	3.8	10.2	5	7.6	6	89	42	47
12	7.2	3	3.0	7.6	0	7.6	6	101	53	48
13	5.4	2	3.0	6 0	5	5.6	0	102	44	58
14	5.6	0	3.0	5.2	0	7.2	0	100	47	53
15	6.8	0	2.2	6.2	0	7.2	0	100	49	51
16	6.6	1	2.4	6.0	2	6.8	2	87	48	39
17	5.8	1	2.6	6.0	0	6.4	2	91	47	44

D, least perceptible difference in grams.
P, per cent. of data showing no discrimination.
MV, statistical mean variation.
B, least perceptible difference for boys.
PB, per cent. of boys showing no discrimination.
G, least perceptible difference for girls.
PG, per cent. of girls showing no discrimination.
N, number of children.
NB, number of boys.
NG, number of girls.

mean variation. When discriminative ability decreases, between any two successive ages, the variation decreases for the corresponding period. On the whole, however, variation decreases with advance in age. At the age of 7, where a falling off in sensitiveness is indicated in the main curve for discrimination, the mean variation is decreased. Also during the years from 12 to 14 the variation is stationary, these too, being years during which the child lost in his ability to discriminate, as shown by chart I. Apparently during these years development has been arrested and consequently those influences removed which would cause variation in the results of different children. At 15, when the discrimination was relatively poor, the mean variation was very small.

In such points the mean variation throws some light upon those years where divergences and abrupt changes occur. Marked changes in the curve for variation would indicate unusual heterogeneity in data at that point. Such results unquestionably represent changes in growth.

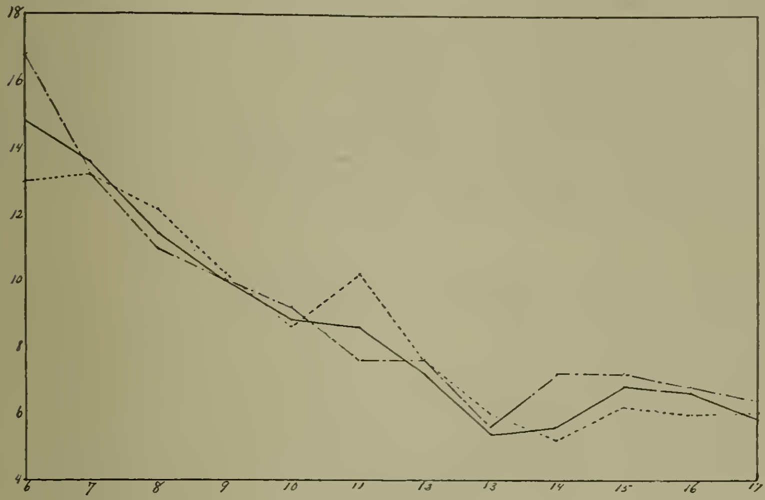

FIG. 6. CHART I.
——— Boys and girls.
-------- Boys.
—-—-— Girls.

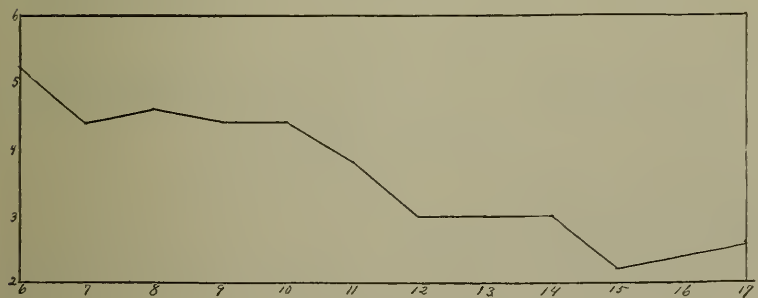

FIG. 7. CHART II.
Statistical mean variation.

TEST (2): *Sensitiveness to color-differences.*

The results of this test are recorded in table II and charts III and IV. The same general rules apply here as in the results of the previous test. The two columns D and P of table II have to be con-

sidered in conjunction, because the scale of shades was insufficient to admit of discrimination of difference in color by all children. As in the previous test the figures in column *P*, *PB* and *PG* indicate the per cent. of children of the respective ages, who said all the colors were exactly alike, discriminating no difference in shade. The numbers on the left, by which the solid line is to be interpreted, indicate the number of colors picked out as being exactly alike. The ages are indicated at the bottom of the chart on the axis of abscissas. Ability to distinguish different shades of the same color increases with age. As a rule, at 7 marked irregularities occur in all the curves which require mental action or discrimination. These irregularities will be spoken of more fully later on.

Table II.

Sensitiveness to color-differences.

Age.	*D*	*P*	*MV*	*B*	*PB*	*G*	*PG*	*N*	*NB*	*NG*
6	9.6	57	1.8	8.3	51	9.6	62	90	45	45
7	9.0	49	2.1	8.3	48	9.6	50	94	50	44
8	8.3	44	2.3	9.6	51	7.0	39	90	44	46
9	6.3	23	2.2	6.1	24	6.6	22	95	45	50
10	5.4	11	1.9	6.0	16	5.2	5	91	48	43
11	5.4	4	1.7	6.0	9	4.9	0	89	41	48
12	5.1	3	1.5	4.8	2	5.1	4	101	53	48
13	4.6	4	1.7	5.2	9	4.1	0	102	44	58
14	4.7	3	1.4	4.8	7	4.6	0	101	47	54
15	4.4	1	1.1	4.1	0	4.6	2	100	49	51
16	4.3	1	1.3	4.3	0	4.0	2	87	48	39
17	3.9	3	1.4	4.0	6	4.9	1	91	47	44

B, least perceptible difference in color in number of shades.
P, per cent. of data showing no discrimination.
M, statistical mean variation.
B, least perceptible difference for boys.
PB, per cent. of data of boys showing no discrimination.
G, least perceptible difference for girls.
PG, per cent of data of girls showing no discrimination.
N, number of children.
NB, number of boys.
NG, number of girls.

In this test the advantage is slightly in favor of the girls. The curves cross and re-cross so frequently, however, that no very plain statement as to comparison of sexes can be given. The boys start at 6 with the advantage of the girls, but at 17 the girls take the lead. By making a general average of all ages for all the boys and all the girls, the advantage of girls over boys is only one-tenth of the difference between the successive shades. Yet, girls have the additional

advantage in that only 18.7% of the girls failed to discriminate at all, while 22.3 % of the boys failed in so doing. This throws the final balance somewhat in favor of the girls. The curve of this sense,

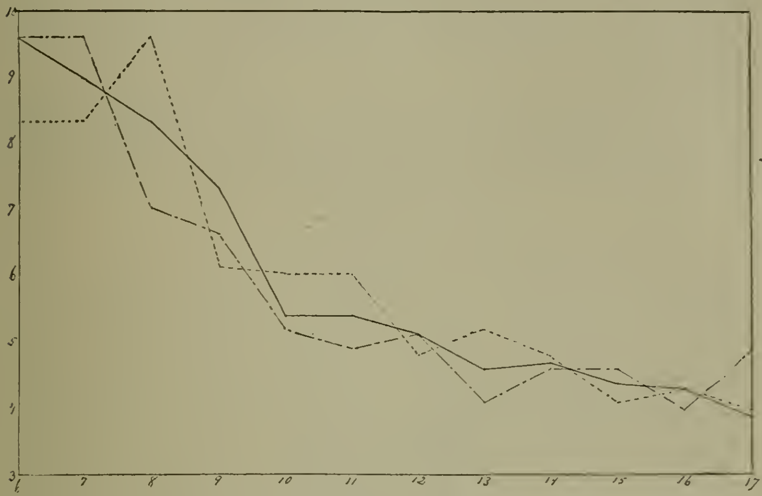

FIG. 8. CHART III.
——— Boys and girls.
.... Boys.
—·—·— Girls.

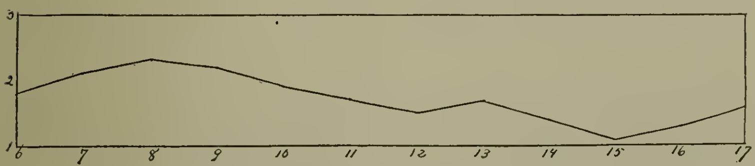

FIG. 9. CHART IV.
Statistical mean variation.

chart II, shows the most gradual *increase* in discriminative ability of any worked out and it will also be noted that the same general regularity in *decrease* of variation is shown in chart IV, with a slight divergence at 13, due probably to puberty.

TEST (3): *Force of suggestion.*

We are continually translating sensations gained from one sense into terms of another sense. In walking, or reaching for articles in the dark, we always imagine how things ought to look and then translate these ideas of sight into muscle sensations in guiding our

muscles. In reaching for the door-knob with closed eyes one always guides his hand by translating how the extended hand looks into how it should feel. The experiments of this test were taken with a view to measuring the influence of the interpretation given by one sense on the decision of another sense, the result being expressed as a function of the age.

The results obtained are recorded in table III and charts V and VI. In considering the results, columns *D* and *P*, *B* and *PB*, *G* and *PG* of table II have to be taken together as in the two preced-

TABLE III.

Force of suggestion.

Age.	*H*	*P*	*MV*	*B*	*PB*	*G*	*PG*	*N*	*NB*	*NG*
6	42.0	36	17.0	43.5	37	42.5	36	92	45	47
7	45.0	37	15.5	43.5	35	43.5	39	95	50	45
8	47.5	27	13.5	45.0	27	49.5	36	92	46	46
9	50.0	36	10.5	50.0	38	49.5	35	94	47	47
10	43.5	23	12.5	40.0	18	44.0	27	91	49	42
11	40.0	22	11.5	38.5	11	40.0	14	91	43	48
12	40.5	15	9.0	38.0	12	41.0	18	103	54	49
13	38.0	8	9.0	37.0	8	38.0	9	103	45	58
14	34.5	7	9.5	31.0	8	33.5	2	100	47	53
15	35.0	12	10.5	33.0	2	38.0	20	100	49	51
16	34.5	6	10.0	32.0	5	38.5	7	86	47	39
17	27.0	5	12.0	25.0	1	31.0	10	84	43	41

H, force of suggestion, in grams.
P, per cent. of data in which the force of suggestion exceeded 65 grams.
MV, statistical mean variation.
B, force of suggestion for boys, in grams.
PB. per cent. of data for boys in which the force of suggestion exceeds 65 grams.
G, force of suggestion for girls, in grams.
PG, per cent. of data for girls in which the force of suggestion exceeded 65 grams.
N, number of children.
NB, number of boys.
NG, number of girls.

ing tests. The figures of columns *P*, *PB* and *PG* of table III indicate the per cent. of data in which the extremes of the series of fourteen blocks were picked out as of the same weight as the respective standards to be compared. The figures at the left of the chart indicate in grams the amount of error made in estimating the differences in weight between the large and small blocks. Ages are marked at the bottom of the chart. As explained under apparatus for test (3), the large and small blocks were both exactly alike in weight, but, owing to the difference in size, the child's judgment as to what the

blocks should weigh by muscle-sense was so influenced by the suggestion from the eye as to what their relative weight should be if judged from sight, that e. g. at 6 they thought there was a difference of 42g between them. In addition to this, in 30% of the data more difference was made between them than could be measured by the limits of my

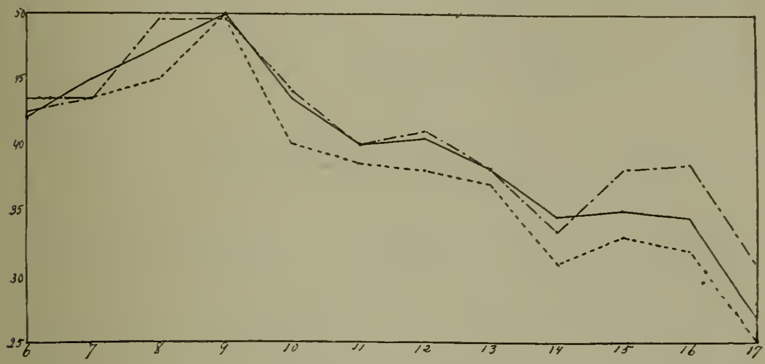

FIG. 10. CHART V.

——— Boys and girls.
--------- Boys.
—-—-— Girls.

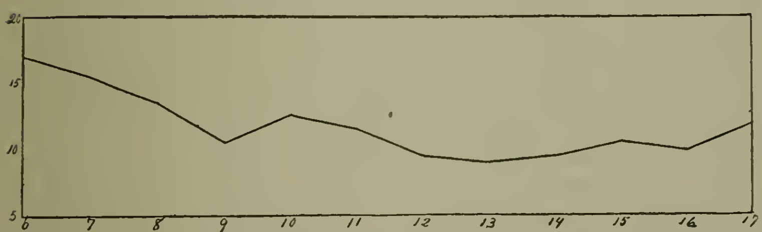

FIG. 11. CHART VI.
Statistical mean variation.

fourteen weights, viz: 65g. At 7 they were influenced by the suggestion of sight even more than at 6. At 7 they made a difference of 45g between the blocks while 37% said there was more difference in their respective weights than my weights would measure, viz: 65g. The influence of the suggestion gradually increased, reaching its maximum at 9 where the average child thought there was a difference of 50g which is almost as much as the weight of the blocks themselves, viz: 55g. In addition to this, at 7, still 36% judged the difference larger than 65g which was the limit of my test. From 9 to 17 this influence gradually decreased, the muscle-sense gradually

learning to correct the suggestion given by sight as to what the relative weight should be. At 17 the large difference of 27g in weight was made between the two blocks while still 5% were found whose error of judgment fell beyond the limits of measurement. As seen in columns *B*, *PB*, *G* and *PG* of table III and also as seen in curves on chart V a marked difference may be noted between boys and girls. Boys, being influenced more by the suggestion, are slightly worse at 6 than girls ; at 7 both are equal ; but thereafter girls are considerably worse than boys with one exception at 9, where the girls and boys may again be said to be the same since the difference is only one-half a gram. The deflection of the curve at age 14 from its general trend is again noticeable, that of the girls being most marked. It is to be presumed that the child grows worse from 6 to 9 because at 6 he has not yet learned to compare, and that as he learns gradually to judge of a thing from more aspects than one, or, in other words, learns to interpret one sense by another, the force of the suggestion given by the eye to the muscle increases until at 9 he has come to the age of experience enough to see that things are not always what they seem. Consequently at this age he begins to correct misleading influences bearing upon him. This error can never be wholly eliminated, for in all my experiments on old as well as young, I have found no one who was not subject to the illusion. The small one was universally chosen as the heavier of the two and not infrequently was it judged to be more than twice as heavy, even by adults. All those who judged that there was more difference in weight between the two blocks than 65g—the limit of my test—it will be easily seen, made the smaller one more than five times as heavy as the larger one. Reference to column *P*, table III, age 7, shows that at that age 37% gave the judgment that the large one weighed 15g or less while the small one weighed 80g or more. Since 15g and 80g were the lightest and heaviest blocks respectively in my series of fourteen, and, since so many picked out these two extremes, it is highly probable that quite a number would have made a much greater difference than 65g between the weights.

The blocks, of course, are seen before being lifted and immediately upon seeing them one judges by sight that the larger one ought to be much the heavier. However, upon lifting and receiving about the same sensation in weight from both we are immediately led to reverse our decision and judge the smaller one to be the heavier.

On the whole, variation decreases with advance in age. In the main curve, chart V, the child becomes worse in his judgment from 6 to 9. In variation, chart VI, he becomes better. In the main curve at 9 he becomes better. In variation curve he becomes worse. At ten the variation again becomes subject to the more general law, however, of decrease with age. The particular law, however, that for short periods in the development where ability increases variation increases, is substantiated by a large proportion of cases in each curve which represent a larger proportion of mental activity.

Test (4): *Voluntary motor ability.*

In column T of table IV are recorded the number of taps the average child can make in five seconds for the respective ages. The results are given in graphic form in charts VII and VIII. The ages are marked along the axis of abscissas, and the figures to the left along the axis of ordinates represent the number of taps made in five seconds. *B* and *G* of the same table are the averages for boys and girls respectively. The average child at 6 years taps 20.8 times in five seconds. From 6 there is a gradual increase until the age 12, reaching at that age a rapidity of 29.9 taps in five seconds. At 13, however, this is lowered to 28.9 taps in five seconds. From this there is the gradual increase again, reaching the maximum at 17 with a rate of tapping amounting to 33.8 taps in five seconds. The data, when calculated for boys and girls separately, show throughout a higher rate of tapping for boys than for girls. Boys at 6 tap 21 times while girls tap only 19.7 times in the five seconds. As the increase in ability goes on, boys always excel in about the same proportion except at 14 and 17 where the difference is much more apparent.[1] Both fall off considerably from 12 to 13. At thirteen however, the boys regain their lost footing and begin to increase again as rapidly as they did before 12. The girls, however, continue to lose until 14 before beginning to gain again. From 16 to 17 they fall off once more. The divergence in this as well as the preceding curves at the period from 12 to 14 is undoubtedly due to the effects of puberty. This would contradict somewhat the statement of Burnham[2] who says that at puberty there is a great increase of

---

[1] Bryan, *On the development of voluntary motor ability,* Am. Jour. Psych., 1893 V 173.

[2] Burnham, *The study of adolescence,* Ped. Sem., 1892 I 181.

vitality and energy and also greater mental activity. The former is undoubtedly true but whether the latter is a justifiable conclusion therefrom is very doubtful. My curves throughout rather seem to justify the opinion of LANGE[1] that physical development takes up the strength and thus retards the mental development.

TABLE IV.

Voluntary motor ability.

Age.	*T*	*MV*	*B*	*MV′*	*G*	*MV″*	*N*	*NB*	*NG*
6	20.8	2.4	21.0	2.5	19.7	2.5	98	49	49
7	22.5	2.9	22.8	2.7	21.2	2.5	98	50	48
8	24.4	2.9	24.9	3.4	23.9	2.2	96	49	47
9	25.4	2.5	25.8	2.5	25.0	2.9	99	50	49
10	27.0	2.8	27.7	2.6	26.9	2.8	97	50	47
11	29.0	3.3	29.7	3.2	27.8	3.0	101	50	51
12	29.9	3.3	30.3	3.1	29.6	3.0	106	56	50
13	28.9	2.8	29.8	3.0	28.1	3.3	110	59	51
14	30.0	3.6	31.2	3.2	28.0	3.4	104	50	54
15	31.1	3.0	31.3	2.6	29.8	3.2	101	51	50
16	32.1	3.3	33.0	3.0	31.8	3.4	87	48	39
17	33.8	2.9	35.0	2.4	31.5	2.3	91	47	44

T, number of taps in five seconds.
MV, statistical mean variation.
B, number of taps for boys.
MV′, statistical mean variation for boys.
G, number of taps for girls.
MV″, statistical mean variation for girls.
N, number of children.
NB, number of boys.
NG, number of girls.

However that may be, for some cause or other, the children must have labored under some disadvantage in almost all my tests at the period about 13. BRYAN's children[2] labored under some similar difficulty at about 13. The individual rate of tapping varied from 14 taps in five seconds by a couple of children 6 years old to 45 taps in five seconds by a boy 17 years of age.[3] The average variations for each age can be seen by referring to table IV and chart VIII.

[1] LANGE, *Uber eine haufig vorkommende Ursache von der langsamen und mangelhaften geistigen Entwicklung der Kinder*, Zt. Psych. Phys. Sinn., 1893 VII 95.

[2] Am. Jour. Psych. 1893 V 204 charts I to V.

[3] No set amplitude of movement was given, thus allowing the child to choose that best adapted to rapidity for himself. Amplitude of movement, however, makes no special difference in rapidity; cf. BRYAN, *Voluntary motor ability*, Am. Jour. Psych., 1893 V 150, 176.

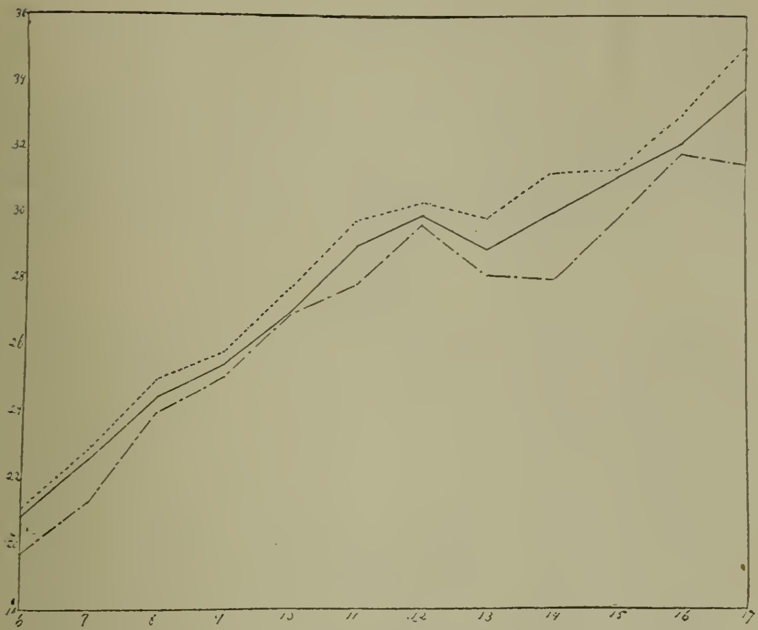

FIG. 12. CHART VII.
——— Boys and girls.
.... Boys.
—-—-— Girls.

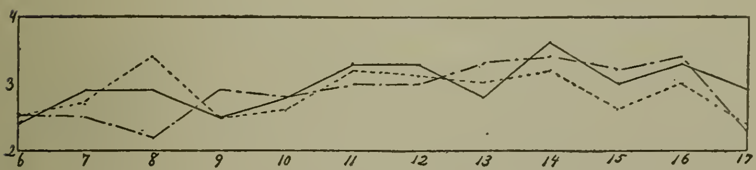

FIG. 13. CHART VIII.
——— Statistical mean variation.
.... Statistical mean variation for boys.
—-—-— Statistical mean variation for girls.

The mean variations for the total result of boys and girls combined were calculated as well as for boys and girls separately, and are found recorded in columns MV, MV' and MV'' respectively. A graphic presentation of the same is also given in chart VIII.

Test (5): *Fatigue.*

After tapping for 45 seconds fatigue entered into the results very noticeably. Column F of table V gives the per cent. of loss between the rapidity of tapping for the first 5 and that for the last 5 of 45 seconds. Columns B and G give the same calculation for boys and girls respectively. MV denotes the amount of deviation of each result from the general average while MV' and MV'' indicate the same for boys and girls respectively. The same results are to be

Table V.

Fatigue.

Age	*F*	*MV*	*B*	*MV'*	*G*	*MV''*	*N*	*NB*	*NG*
6	21.4	8.1	22.8	9.4	21.3	7.0	98	49	49
7	21.0	8.9	22.5	9.7	20.2	6.7	98	50	48
8	24.0	7.3	24.7	8.3	23.3	7.1	96	49	47
9	21.0	7.1	22.5	6.7	20.7	7.8	99	50	49
10	22.0	7.5	22.7	7.8	19.0	7.1	97	50	47
11	20.0	6.2	20.3	6.5	18.0	5.5	101	50	51
12	16.0	6.3	18.0	6.0	14.0	6.7	106	56	50
13	14.5	6.4	15.8	6.7	14.7	5.8	110	59	51
14	14.0	6.5	17.8	6.2	12.0	6.1	104	50	54
15	12.7	5.8	13.8	4.9	11.5	5.7	101	51	50
16	14.7	5.2	15.3	4.6	11.7	5.6	87	48	39
17	13.8	5.3	14.5	6.3	13.5	4.3	91	47	44

F, per cent. of loss in rapidity of tapping after tapping 45 seconds.
MV, statistical mean variation.
B, per cent. of loss in rapidity for boys.
MV', statistical mean variation for boys.
G, per cent. of loss in rapidity for girls.
MV'', statistical mean variation for girls.
N, number of children.
NB, number of boys.
NG, number of girls.

found in graphic form in charts IX and X. Ages are marked at the bottom. The figures to the left of the chart indicate the per cent. of loss in rapidity of tapping between that of the first five and that of the last five seconds. The average child at 6 loses 21.4% after tapping 45 seconds. From 6 to 7 a slight gain is made, the loss by fatigue being 21% at 7. At 8, however, the effect of fatigue is much more marked, this being the age at which the child loses most rapidly; here there was a loss of 24%. After 8 the fatigue is less and less noticeable till the age of 15 where it was least marked, being only 12.7%. From 15 to 16 it again becomes more marked, a loss of 14.7% occurring at 16 with a succeeding gain again at 17, where it was 13.8%.

When these data are calculated for boys and girls separately, it becomes evident that girls tire more easily at 13 than at 12 while

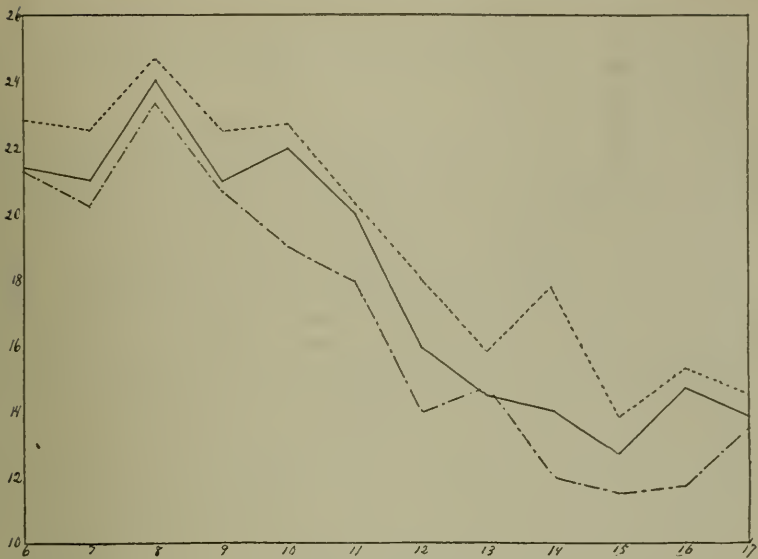

FIG. 14. CHART IX.
——— Boys and girls.
.... Boys.
—·—·— Girls.

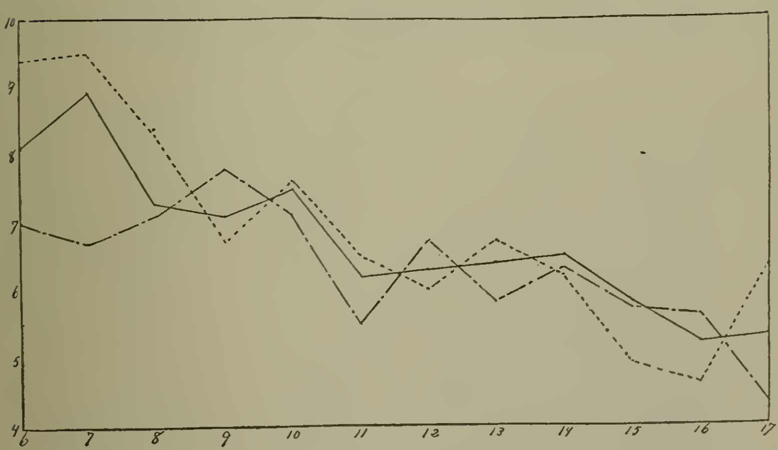

FIG. 15. CHART X.
——— Statistical mean variation.
.... Statistical mean variation for boys.
—·—·— Statistical mean variation for girls.

for boys this variance comes a year later between 13 and 14. As in almost all of the charts representing mental research, there seems

to be a marked turn in the life of the child at 7. This divergence is brought out very plainly also by the mean variations shown in chart X for ages 6 and 7.

Boys tire more *quickly* throughout in voluntary movement than girls. But the statement that boys tire more *easily* than girls could scarcely be made upon the basis of my data for it will be remembered that the rate of tapping by the boys, as shown by table IV and chart VII, was faster than that by the girls. The statement that boys tire more *easily* is unwarrantable, for, by averaging and comparing the rate of tapping for all boys and girls separately, it is found that the girls on the whole tap slower than the boys who lose but little more than the girls by fatigue, leaving the balance in favor of boys. The average boy, including all ages, taps 29.4 times in five seconds, the average girl taps 26.9 times, thus tapping 8.5% slower than boys. The average boy, including all ages, loses 18.1% by fatigue; the average girl loses 16.6%. In other words, the boys lose 1.5% more by fatigue than girls and yet boys tap 8.5% faster than girls. This leaves the balance greatly in favor of boys when voluntary motor ability and fatigue are considered together.

Test (6): *Weight.*

Column *W* of table VI indicates the weight in pounds according

Table VI.

Weight.

Age.	*W*	*MV*	*B*	*MV′*	*G*	*MV″*	*A*	*B*	*C*	*N*	*NB*	*NG*
6	46.0	4.6	46.8	4.4	44.3	4.3	43.3	46.8	49.0	98	50	48
7	51.0	4.8	51.2	4.7	50.4	4.4	48.3	51.0	51.5	98	50	48
8	53.0	5.9	52.5	6.0	53.0	5.1	53.3	53.5	53.3	96	49	47
9	59.5	6.2	60.0	9.9	58.8	6.8	59.0	59.0	61.5	97	49	48
10	66.5	7.7	68.4	6.9	62.7	7.4	67.2	64.5	66.8	96	50	46
11	70.0	7.8	70.8	6.7	70.0	6.0	70.0	70.0	66.5	101	51	50
12	83.5	12.3	82.3	6.7	84.5	11.5	83.0	83.3	87.8	106	56	50
13	89.5	11.6	88.0	9.4	92.0	10.6	82.2	92.3	86.0	110	59	51
14	96.0	15.4	91.7	15.8	98.0	13.3	98.5	98.8	90.3	104	50	54
15	105.0	13.7	110.0	15.4	104.0	10.5	105.5	106.0	105.0	102	51	51
16	119.8	15.4	127.0	11.9	113.0	11.7	116.0	124.5	118.3	87	48	39
17	122.0	14.9	130.0	11.3	113.7	15.1	120.0	122.8	125.0	90	46	44

W, weight in pounds.
MV, statistical mean variation.
B, weight of boys.
MV′, statistical mean variation for boys.
G, weight of girls.
MV″, statistical mean variation for girls.

A, weight of bright children.
B, weight of average children.
C, weight of dull children.
N, number of children.
NB, number of boys.
NG, number of girls.

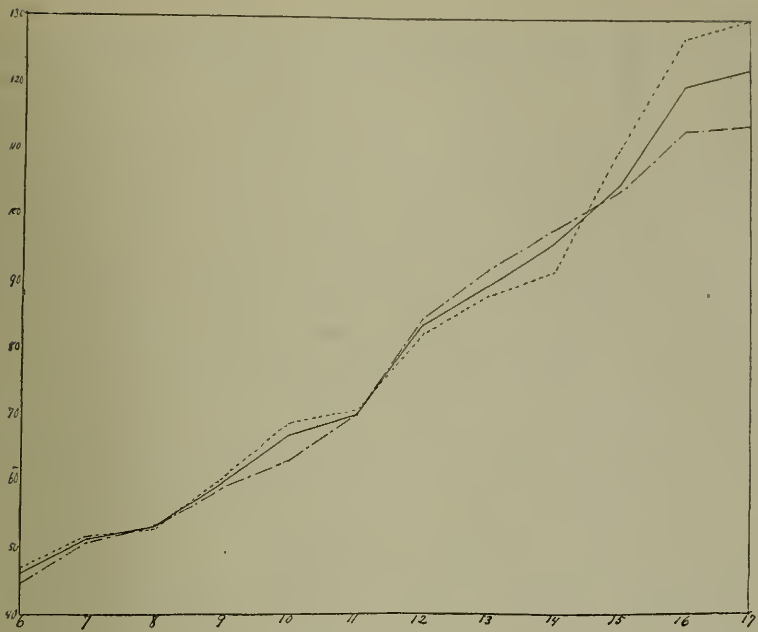

FIG. 16. CHART XI.
——— Boys and girls.
.... Boys.
—-—-— Girls.

FIG. 17. CHART XII.
——— Statistical mean variations.
.... Statistical mean variation for boys.
—-—-— Statistical mean variation for girls.

to age. The individual weights were taken in quarter-pounds. The weights of boys and girls separately are to be found in columns *B* and *G*. In chart XI the ages are at the bottom and the weight in pounds to the left of the chart. The figures at the left of chart XII indicate the variation in pounds. The average weight at 6 years was 46 pounds; this in general increases with advance in age, the weight at 17 being 122 pounds. Certain differences are noticeable in the relative rapidity of growth between different ages. Boys have their most rapid growth between 14 and 16, increasing in weight 18.3 pounds between 14 and 15, and 17 pounds between 15 and 16, but between 16 and 17 the increase in weight is very slight indeed, being only 3 pounds. The most rapid growth for girls occurs between 11 and 12, being 14.5 pounds. Up to the age 12 boys and girls seem to grow in about the same proportions, boys being slightly heavier than girls. Between 11 and 12 the order is reversed, girls growing faster and becoming heavier than boys ; they remain heavier until between 14 and 15. Between 14 and 15 boys again begin very rapid growth and from then on are much heavier than girls.

At the age 11, as shown in chart XI, the girls begin the period of most rapid growth. Chart XII, showing the mean variation for weight, indicates also a sudden rise in the mean variation at that time. In weight the mean variation increases with advance in years. The contrary was true in the three preceding curves where mental work was involved. It will be remembered also, that in the preceding curves, wherever there was a sudden decrease between two successive ages in ability to discriminate, there was a sudden decrease in the mean variation for the corresponding period. In these purely physiological data, however, the opposite seems to be true. The mean variation in chart XII rises at the point corresponding to the one in chart XI where the rapid growth of the girls begins. Boys in chart XI begin their rapid growth later than girls and in chart XII the sudden rise in the mean variation for the boys begins a year later. From the comparison for "bright," "average" and "dull" the same negative conclusion is to be drawn as in the following test.

Test (7): *Height.*

The height was taken in tenths of a centimeter. Column *H* of table VII is the record of height, the upper figures of each age being in inches, the lower in centimeters and tenths. The height of boys and girls separately is to be found under *B* and *G* respectively.

The average individual mean variations for all combined and for boys and girls separately are expressed in columns *MV*, *MV'* and *MV''* respectively. The figures on the left of chart XIII indicate the height in centimeters. Those at the left of chart XIV indicate the variations in centimeters. The ages are marked below in both charts. Almost precisely the same laws appear here in regard to rapidity of growth for the different sexes as appeared in the figures for weight. At 6 the boys are 114.5^{cm} high, the girls 114.0^{cm}. Both boys and girls grow with about the same rapidity, the boys being the taller, until between 11 and 12; here the girls grow much more rapidly and are the taller until between 14 and 15, where the boys are taller. The girls become more nearly stationary in height after 15, while boys make exceedingly rapid progress from 14 on. At 17 the height of boys was 170.5^{cm}, of girls 168.6^{cm}. Just before puberty is the period of most rapid growth for girls while the period of most rapid growth falls later for boys, beginning at 14.

The statement has been made by Porter that the brighter the child the taller he is.[1] Brightness and dullness, however, in his tests were decided by examination-grades, which, it is needless to say, are often very poor mental tests. In my results no such relation could be traced. My data are based upon the judgment of the teacher as to what she considered to be the general mental ability or "stand" of the child as it is sometimes called. This is really more accurate than a system of set examinations. The results as tabulated in columns *A*, *B* and *C* of table VII, represent the heights of the children graded according to the judgment of the teacher under whom they fell, *A* being bright, *B* those of average ability, and *C* those who were dull mentally. In the same results, when put in graphic form, the lines cross and re-cross too frequently to be of any value on such a point except to give the negative result as disproof of the statement referred to above.

The mean variations in chart XIV furnish one marked exception in the curve for boys to the rule applying so well to weight and even still more forcibly corroborated by the next curve of variation for lung capacity, chart XVI. Chart XIV for variations in height is not without points verifying the rule, however. Variations increase in size with advance in age. Following the curve for girls, at 10 and 11, where the period of very rapid growth begins to show

---

[1] Porter, *The growth of St. Louis children*, Transactions of the Academy of Science of St. Louis, 1894 VI 335.

TABLE VII.

Height.

Age.	*H*	*MV*	*B*	*MV'*	*G*	*MV''*	*A*	*B*	*C*	*N*	*NB*	*NG*
	45.4	1.6	45.0	1.6	44.9	1.4	44.2	45.2	45.6			
6	115.4	3.9	114.5	3.9	114.0	3.6	112.3	114.8	115.8	98	50	48
	47.3	1.6	47.1	1.6	46.9	1.5	47.4	46.8	47.6			
7	120.2	4.0	119.8	4.0	119.1	3.5	120.5	119.0	121.0	98	50	48
	49.3	2.0	48.9	1.8	48.4	1.8	48.5	48.4	48.2			
8	125.2	4 9	124.2	4.4	123.0	4.4	123.3	122.8	122.3	96	49	47
	51.2	2.1	51.2	2.0	50.8	1.8	51.0	50.9	50.9			
9	130.2	5.3	130.2	5.1	129.0	4.5	129.4	129.2	129.3	97	49	48
	53.0	2.0	53.0	1.7	52.8	2.1	52.9	53.0	53.0			
10	134.6	4.9	134.6	4.3	134.0	5.2	134.4	134.6	134.5	96	50	46
	55.9	2.4	55.9	2.0	54.6	2.2	55.1	54.9	55.9			
11	142.0	6.0	142.0	5 0	138.6	5.6	140.0	139.4	142.7	101	51	50
	57 3	2.4	57.0	2.2	57.9	2.5	57.6	57.3	59.0			
12	145.5	6.1	144.8	5.6	147.1	6.3	146.2	145.5	150.0	106	56	50
	59.9	2.6	58.8	2.2	60.4	2.3	58.0	60.5	58.8			
13	151.0	6 4	149.4	5.6	153.4	5.8	147.2	153.8	149.3	110	51	59
	60.9	3.2	59.3	3.4	61.4	2.8	61.1	60.7	60.4			
14	154.7	8.1	150.5	8.7	155.9	7.1	155.2	154.2	153.3	104	50	54
	62.7	2.9	62.8	3.2	62.5	2.1	64.5	62.6	62.2			
15	159.2	7.4	159.5	8.0	158.8	5.3	163.8	159.0	158.0	102	51	51
	64.9	2.4	65.7	2.2	62.5	2.1	65 0	65.5	65.0			
16	164.8	6.0	167.0	5.4	158.8	5.2	165.1	166.3	165.2	87	48	39
	65.6	2.3	67.1	1.3	63.6	1.9	66.5	65.6	65.2			
17	166.6	5.9	170.5	3.2	161.6	4.7	169.0	166.6	165.6	91	47	44

Upper figures are inches; lower figures are centimeters.
H, height.
MV, statistical mean variation for total result.
B, height of boys.
MV', statistical mean variation for boys.
G, height of girls.
MV'', statistical mean variation for girls.
A, height of bright children.
B, height of average children.
C, height of dull children.
N, number of children.
NB, number of boys.
NG, number of girls.

itself, the variation rises rapidly. At 12, after the most rapid section of increased rate of growth is completed, the variation falls. With the exception of the one point at 14, chart XIV, which is probably due to puberty, the curve of variations gradually decreases with the corresponding decrease in rapidity of growth, shown in chart XIII. The curve for boys contradicts the rule. There is no evident

cause for the sudden fall in variation from 9 to 10 and again at 14. The curve, instead of rising with the rapid growth shown in chart XIII, falls rapidly till the age 17.

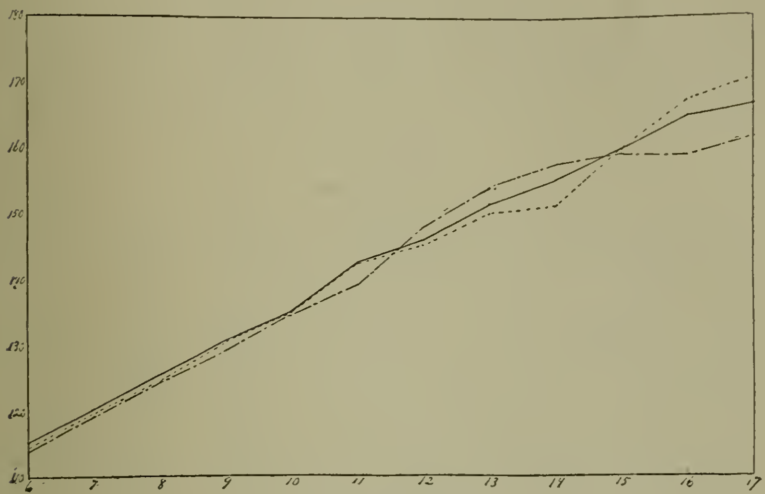

FIG. 18. CHART XIII.
——— Boys and girls.
.... Boys.
— - — — Girls.

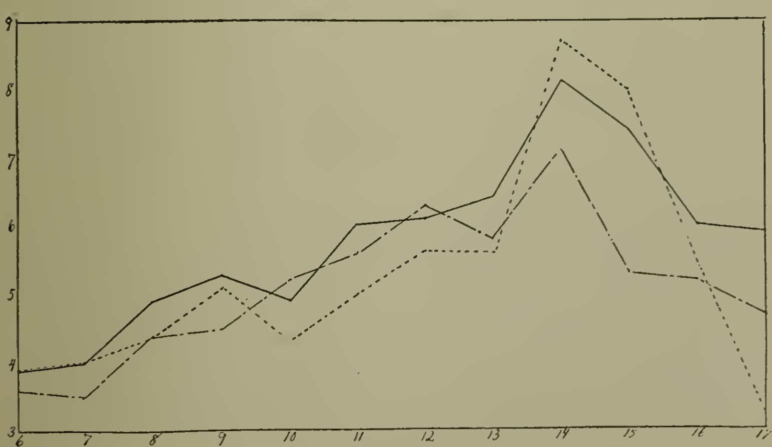

FIG. 19. CHART XIV.
——— Statistical mean variation.
.... Statistical mean variations for boys.
— - — — Statistical mean variations for girls.

Test (8): *Lung-capacity.*

The figures in column *LC* of table VIII indicate the lung-capacity for the corresponding ages given in the first column. The upper figures of each age represent the number of cubic inches; the lower figures represent the number of cubic centimeters. These results are also placed in graphic form in chart XV, the figures at the left indicating the number of cubic centimeters, those at the bottom indicating the ages.

Table VIII.

Lung-capacity.

Age.	*LC*	*MV*	*B*	*MV'*	*G*	*MV''*	*A*	*B*	*C*	*N*	*NB*	*NG*
	52.0	9.7	56.0	9.5	50.0	9.5	52.3	51.3	49.0			
6	832	155	896	152	800	152	837	821	784	94	49	45
	63.0	11.8	66.0	11.2	54.0	11.1	67.0	60.0	67.0			
7	1008	189	1056	179	864	178	1072	960	1072	96	48	48
	72.0	12.3	73.0	11.3	66.0	10.5	72.5	68.5	67.0			
8	1152	197	1168	181	1056	168	1160	1096	1072	96	49	47
	78.0	13.8	83.0	16.3	72.5	9.8	78.0	72.0	81.0			
9	1248	221	1328	261	1160	157	1248	1152	1296	94	48	46
	87.5	14.3	91.5	15.4	82.0	14.3	90.0	84.0	80.5			
10	1400	229	1464	246	1312	229	1440	1344	1288	96	50	46
	91.0	15.9	104.0	15.1	83.0	10.4	95.0	90.0	91.0			
11	1456	254	1664	242	1328	166	1520	1440	1456	100	49	51
	109.0	16.6	113.5	14.1	104.0	14.8	108.0	110.0	113.0			
12	1744	266	1816	226	1664	237	1728	1760	1808	103	56	47
	115.0	19.4	120.0	17.5	105.0	18.8	116.0	116.0	102.0			
13	1840	292	1920	280	1680	301	1856	1856	1632	107	50	57
	117.5	22.7	125.0	23.9	105.0	17.4	119.5	118.5	97.5			
14	1880	363	2000	382	1680	278	1912	1896	1560	101	49	52
	131.3	22.3	161.0	29.8	116.0	15.5	143.5	123.0	131.0			
15	2101	356	2576	477	1856	248	2296	1968	2096	102	51	51
	149.0	39.0	187.0	30.8	115.0	16.1	137.0	166.3	128.5			
16	2384	624	2992	493	1840	258	2192	2660	2056	87	48	39
	156.0	42.0	204.0	33.4	118.5	18.5	180.0	129.5	169.0			
17	2496	672	3264	534	1896	296	2880	2072	2704	91	47	44

Upper figures are cubic inches; lower figures are cubic centimeters.
LC, lung-capacity.
MV, statistical mean variation for total result.
B, lung-capacity for boys.
MV', statistical mean variation for boys.
G, lung-capacity for girls.
MV'', statistical mean variation for girls.
A, lung-capacity of bright children.
B, lung-capacity of average children.
C, lung-capacity of dull children.
N, number of children.
NB, number of boys.
NG, number of girls.

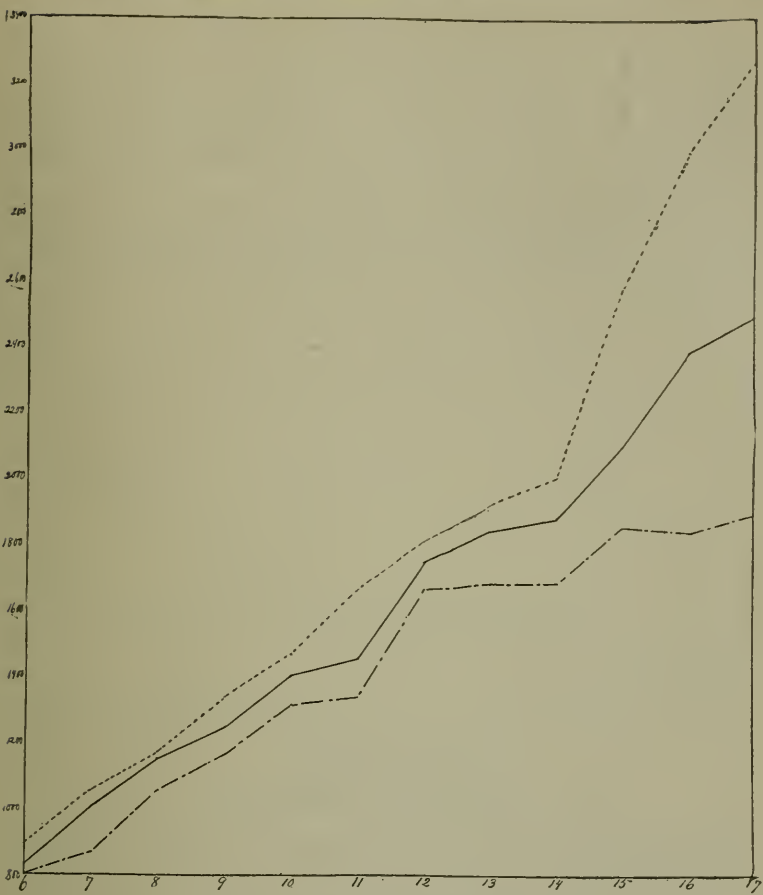

FIG. 20. CHART XV.

——— Boys and girls.
.... Boys.
— - — - — Girls.

Boys have a larger lung-capacity than girls throughout. At 6 boys have a capacity of 896^{ccm} while the girls have only 800^{ccm}. This difference between boys and girls increases but very slightly, in favor of boys, till the age of 12 where boys have a capacity of 1816^{ccm} and girls 1664^{ccm}. From this age on the development of girls is much slower. Almost no growth occurs between 12 and 14. From 14, at which age their lung-capacity is 1680^{ccm}, a gain is made to 1856^{ccm} at 15. From 15 to 17 almost no development of the lungs is noticeable, for at 17 they have attained only 1896^{ccm}. The curve for boys shows a marked difference from that of girls. About the same rate of growth as that from 6 to 7 was maintained till the age of 14, at

which age their lung-capacity was 2000^{ccm}. While the girls become nearly stationary at 12, the boys do not begin their most rapid growth until 14. Between 14 and 15 they make a gain of 576^{ccm}, and they keep up this rapid growth until 17, which is as far as my experiments extended. At 17 the lungs of the average boy contained 3264^{ccm}; those of the average girl 1896^{ccm}.

The mean variations are put in graphic form in chart XVI. The numbers to the left indicate the average mean variation in cubic centimeters.

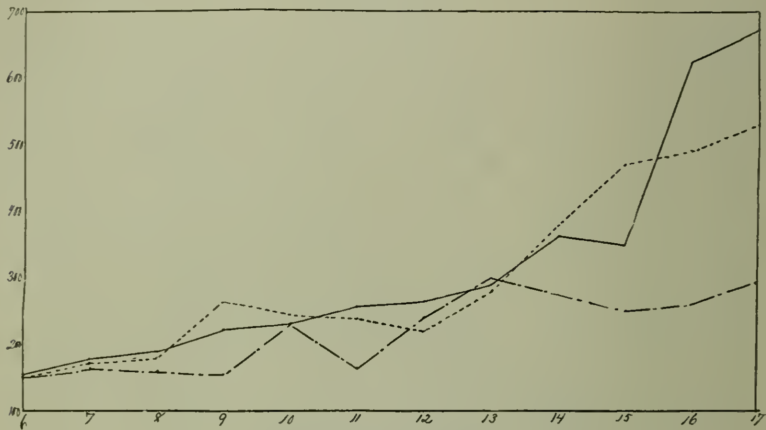

Fig. 21. Chart XVI.

——— Statistical mean variation.
--------- Statistical mean variation for boys.
— - — - — Statistical mean variation for girls.

The mean variation increases with advance in years. From 6 to 13 there is a gradual rise in the variation. At this age the increase in lung-capacity almost ceases for girls and so also does the variation. Chart XV shows almost no increase in lung-capacity between the years 12 and 14. The variation for the corresponding years in chart XVI shows a fall but with a rise again at 14 where the lung-capacity begins to increase as shown by chart XV.

At 13 for boys it is just the opposite. Variations rise very rapidly as does the growth in lung-capacity, the latter appearing one year later. The coincidence between changes in growth and changes in variation are very noticeable in all the physiological curves.

The mean variation, for both boys and girls combined, as is shown in the solid line of chart XVI, increases very rapidly from 13 to 17 owing to the fact that girls after 13 grow but little more while boys undergo most rapid growth thereafter, thus of course throwing the mean variation for the total result much higher.

Test (10): *Reaction-time.*

In taking up this section of the results we return to the mental processes in contra-distinction to the three preceding tests, which were purely physiological. In explaining apparatus and methods, tests (9) and (10) were treated in a different order from that in which they were taken, here it seems convenient and proper also to reverse the order on account of the relative complexity of the two. The simple reaction time alone will be considered first. In the following three tests the arithmetical averages were calculated as well as the median values. In order to show the difference between the results, as calculated by the two different methods I have tabulated the arithmetical averages under *Ta*, and the median values under *Tp*. The same comparison is drawn by the graphic method in the chart, the dotted line being the arithmetical averages and the solid lines the median values. In order to illustrate the method of median values I have selected from my results a card of a boy 8 years old. The point of objection to the arithmetical average is brought out somewhat more plainly in the following figures than would be the case with the average card but by choosing a somewhat extreme case, the underlying principle can be most easily seen and the difference made all the more forcible for illustration. As was explained under methods, each child was given ten trials. The results for the ten reactions of this 8-year-old boy were as follows : 24, 23, 48, 22, 21, 24, 43, 21, 22, 19. The arithmetical average of these ten is 26.7, whereas the median value, or the value half way between the fifth and sixth counting either from the largest or from the smallest as 1, is 22.5. The variations of each result from the mean value are 1.5, 0.5, 25.5, 0.5, 1.5, 1.5, 20.5, 1.5, 0.5, and 3.5 respectively, making a mean variation of 5.7. A glance at the original data will show that in order to get what we would consider the representative reaction-time of that boy, 48 and 43 ought not have as much influence given them as 21, 22 and 23 which fall nearer the average. By the median value, instead of allowing them to count in direct proportion to their size, they are rather allowed to count as one in a series of ten. However, these extreme data such as 48 and 43, are not allowed to pass unnoticed by the method of median values without exciting any influence whatever. On the contrary their influence is felt in the mean variation for the child where their effect properly belongs. In this the average of mean variations for the child is 5.7. Had it not been for the two figures 48 and 43 his mean variation would have been only 1.4 instead of the 5.7. They increase this average and

TABLE IX.

Reaction-time.

Age.	*Ta*	*Tp*	*mv*	*MV*	*B*	*MV'*	*G*	*MV''*	*A*	*B*	*C*	*N*	*NB*	*NG*
6	31.7	29.5	5.6	5.0	28.2	4.6	29.5	5.4	28.5	29.5	28.3	99	50	49
7	30.9	29.2	5.4	5.5	26.7	4.6	31.5	5.2	29.2	29.5	28.2	98	50	48
8	28.7	26.2	4.9	3.9	24.5	3.9	26.0	3.1	26.6	25.8	28.5	96	49	47
9	26.9	25.0	4.1	4.1	24.3	5.4	25.5	4.9	23.2	25.3	26.8	99	50	49
10	23.3	21.5	4.2	3.6	21.0	2.6	22.5	4.3	21.0	20.5	26.0	97	50	47
11	21.0	19.5	3.7	3.4	18.5	3.1	20.6	3.4	19.8	19.8	19.2	101	50	51
12	20.7	18.7	3.6	3.1	17.8	2.7	19.8	3.5	18.3	19.5	19.5	106	56	50
13	20.5	18.7	3.3	3.0	17.8	2.9	20.5	3.5	18.5	18.5	22.5	110	51	59
14	19.1	18.0	3.0	2.9	18.0	3.0	18.7	3.0	16.8	17.6	19.3	104	50	54
15	18.4	17.2	3.0	2.7	16.7	2.3	18.9	2.7	16.0	17.0	19.0	102	51	51
16	17.0	15.5	2.8	2.3	14.7	1.6	17.2	2.6	15.5	16.0	15.0	87	48	39
17	17.0	15.5	3.0	3.3	14.7	1.9	16.3	2.6	14.8	15.5	16.2	91	47	44

Ta, reaction-time in hundredths of a second—arithmetical averages.

Tp, reaction-time in hundredths of a second—median values.

mv, average individual mean variation.

MV, statistical mean variation.

B, reaction-time for boys in hundredths of a second.

MV', statistical mean variation for boys.

G, reaction-time for girls in hundredths of a second.

MV'', statistical mean variation for girls.

A, reaction-time of bright children.

B, reaction-time of average children.

C, reaction-time of dull children.

N, number of children.

NB, number of boys.

NG, number of girls.

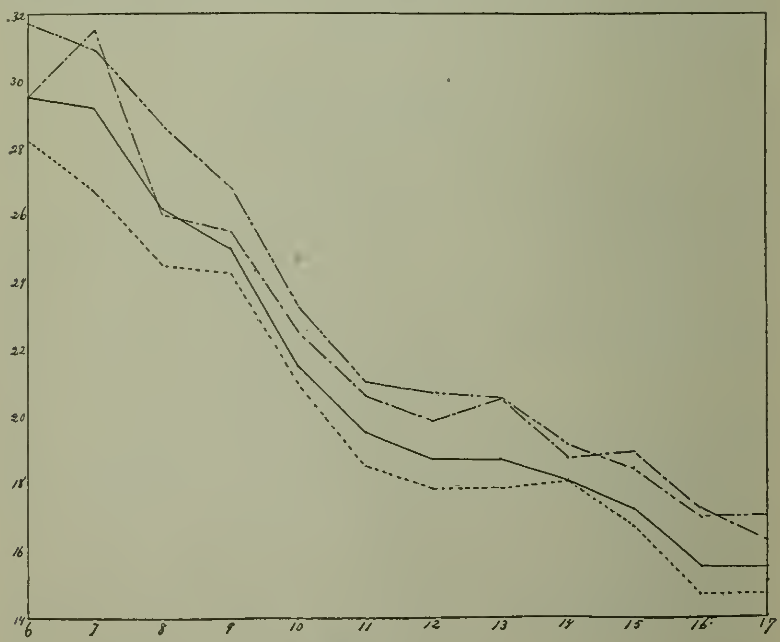

FIG 22. CHART XVII.

———— Boys and girls.
.... Boys.
—·—·—·— Girls.
—··—··— Boys and girls, arithmetical average.

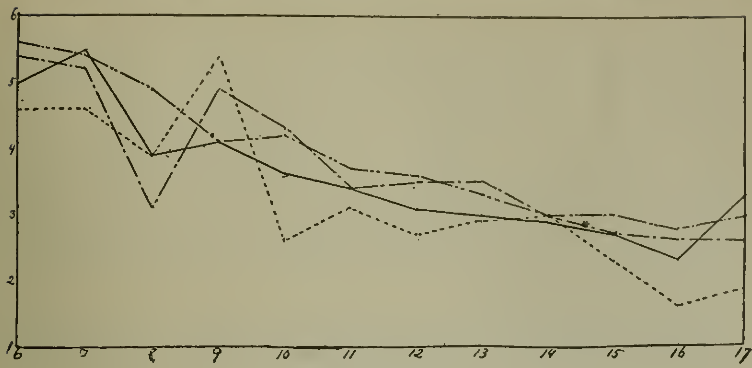

FIG. 23. CHART XVIII.

——————— Statistical mean variation.
.... Statistical mean variation for boys.
—·—·—·— Statistical mean variation.
—······— Average individual mean variation.

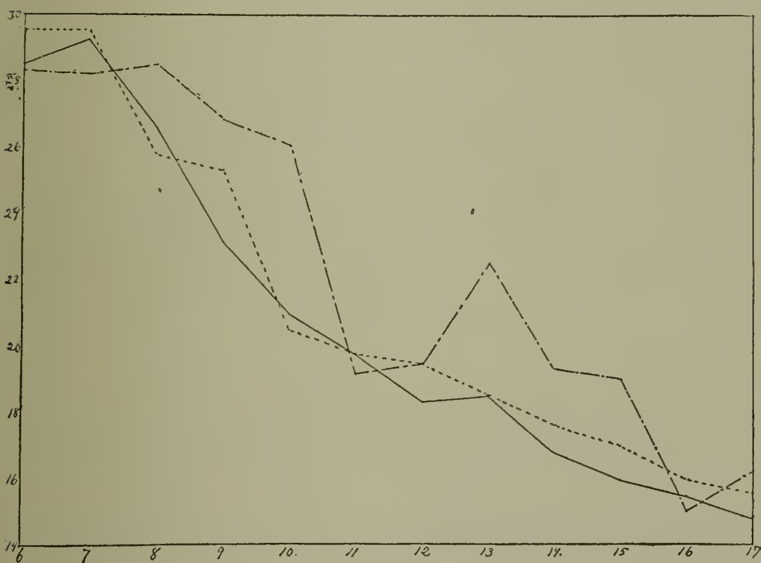

FIG. 24. CHART XIX.

——————— Bright.
.... Average.
—·—·— Dull.

justly so, for, by the mean variation we wish to indicate the irregularity of his separate results from the general average. The median was calculated for both boys and girls combined, and also for each separately. The mean for boys and girls combined according to arithmetical averages always falls half way between those for boys

and girls separately ; not so with the median. The median for boys and girls combined frequently falls below or above both that of the girls and that of boys when taken separately. Illustrations of this can be seen in charts IX, age 13 ; XII, age 7 ; XVII, age 8; XXII, age 14 ; XIII ages 6, 7, 8 and 11. This difference of the median from the average clearly shows the heterogeneity of the two classes.

It must also be noted that mean variations for boys and girls combined are much larger than for each separately, which also shows the heterogeneity of the data.

The time of simple reaction decreases with age. Boys and girls at 6, when averaged together, react in 29.5 hundredths of a second. This decreases to the age 12 where the time is 18.7 hundredths of a second. From 12 to 13 no increase is made, remaining at 18.7 for 13 also. From 13 on, there is gradual increase until 16 when the time is 15.5 hundredths of a second. At 17 no gain is made.

The results, when considered for girls and boys separately, show marked difference in sex. Girls are slower at 7 than at 6. At 6 the time required was 29.5 while at 7 they required 31.5. At 8 there was a gain to 26.0. From this on there was a gradual gain in ability and a decrease in time till 12, where the time was 19.8. At 13, however, 20.5 hundredths of a second were again required leaving the girls only one thousandth of a second better at 13 than they were at 11. After this there was an increase again till 17 where the reaction-time for girls was 16.3.

The curve for boys shows no change from the general law of increase from 6 to 7. From 12 to 14 there is a marked difference in the rapidity of increase. At 12 the time required was 17.8 ; at 13 it was the same ; at 14 there is a loss in ability, the time being 18.0. Thus the boys were worse at 14 than at 12 and but very little better than they were at 11. After 14 they again increased with almost the same rapidity as they did before 11 until 16 and 17 where 14.7 hundredths of a second were required. Both boys and girls seemed to increase less rapidly from eight to nine than at the other ages. Boys were quicker than girls throughout.

The mean variations, represented in chart XVIII, decrease with advance in years. For boys, during the period from 11 to 14 when the actual time of reaction remained almost stationary, the mean variation did the same, showing a slight difference at 12 in the same way as the curve for boys in chart XVII. There is also a marked break in the curve of variation from 8 to 9 corresponding to the slight decrease in the rate of ability at the same ages in chart XVII.

The relation between mean variation and mean reaction-time is rendered most marked by comparing the curves in charts XVII and XVIII for both boys and girls combined. At most points where there is a marked change in rate of increase in chart XVII there is also a corresponding change in the variation at that age. In this, age 13 offers the only exception. Wherever the rate of increase in ability from one age to another is less than the average rate of increase, the variation for that period becomes higher and therefore worse.

In this test also the data were separated and recalculated to find the reaction-time of those who were bright, of average mental ability and dull respectively. The results are recorded in columns *A*, *B* and *C* respectively, table IX. The difference here becomes very noticeable as can be seen by referring to the graphic representation, chart XIX. The bright children react much more quickly than the dull. Not so much difference is noticeable between those who were considered bright by the teacher and those who were judged of average ability. It is shown here that we judge of a child's mental ability by the quickness or rapidity with which they were able to act. Another fact is that all children are considered of about equal mental ability, or in other words, all grades of children react in about the same length of time just before those ages in which changes of growth manifested themselves, viz: 11 and 16. The average reaction-time of all ages for bright children was 20.7 hundredths of a second; for those of average ability it was 21.3; for dull children 22.4.

Test (9): *Reaction with discrimination and choice.*

Here, as in the other mental tests, ability increased and the length of time required decreased with advance in age. This test implies more complicated mental activity and, as would be expected, the influences which affect mental life show themselves more plainly in the curve representing such development. For some cause or other development between 6 and 7 is arrested for girls here, as well as in the test on reaction-time. Boys seem to suffer no such back-set but, starting at 53.5 hundredths of a second, continually increase from 6 till 13. From 13 to 14 they suffer a slight loss after which they gain till 17, losing slightly, however, from 15 to 16. At 17 the time required for boys was 30.5 hundredths of a second. Boys may be said to undergo only one loss, that being also of small moment. Girls suffer two marked losses, the first from 6 to 7, increasing the time required from 51 hundredths to 52.8 hundredths of a second. After 7

TABLE X.

Reaction with discrimination and choice.

Age.	*Ta*	*Tp*	*mv*	*MV*	*E*	*B*	*MV'*	*G*	*MV''*	*A*	*B*	*C*	*N*	*NB*	*NG*
6	55.8	52.5	10.2	6.0	1.1	53.5	5.3	51.0	6.5	52.3	51.0	54.5	99	50	49
7	54.1	53.0	9.4	8.1	1.1	49.0	8.8	52.8	9.4	52.0	52.7	55.5	97	49	48
8	48.8	47.8	8.5	6.5	.9	48.0	5.7	47.5	5.5	47.5	47.5	50.0	96	49	47
9	47.5	45.0	8.1	6.8	1.2	44.5	6.3	46.0	7.2	45.8	45.0	42.3	98	49	49
10	42.2	41.0	7.3	4.9	1.2	40.0	4.9	41.5	4.5	39.7	41.0	45.5	97	50	47
11	40.5	38 5	7.0	5.8	1.2	38.7	5.8	38.8	5.7	39.0	38.0	38.0	101	50	51
12	38.9	37.0	6.1	5.5	1.1	38.5	6.0	37.0	4.9	37.0	36.5	37.0	106	56	50
13	39.9	39.5	6.2	5.8	1.2	36.0	5.1	41.5	5.5	39.0	38.1	42.8	110	51	59
14	36.3	36.5	6.5	4.9	.9	36.7	4.5	35.5	5.4	33.5	36.3	36.8	104	50	54
15	34.8	33.5	5.9	4.9	.8	31.1	5.5	34.5	3.8	30.5	34.0	36.5	102	51	51
16	34.0	32.5	5.4	4.3	1.0	31.5	3.9	35.0	3.9	32.7	32.7	34.3	87	48	39
17	32.1	31.2	5.4	4.0	.7	30.5	3.5	31.5	4.4	32.5	30.0	31.2	91	47	44

Ta, discrimination-time in hundredths of a second—arithmetical averages.

Tp, discrimination-time in hundredths of a second.

mv, average individual mean variation.

MV, statistical mean variation.

E, average number of errors made by reacting to red.

B, discrimination-time for boys in hundredths of a second.

MV'. statistical mean variation for boys.

G, discrimination-time for girls in hundredths of a second.

MV'', statistical mean variation for girls.

A, discrimination-time for bright children.

B, discrimination-time for average children.

C, discrimination-time for dull children.

N, number of children.

NB, number of boys.

NG, number of girls.

they increase in ability very rapidly till the age 12, where the length of time was 37 hundredths. From 12 to 13, however, they lose just as much as they had gained during the two years preceding 12, thus requiring 41.5 hundredths of a second at 13 which is the same length of time required as at age 10. After 13 comes another very rapid gain till 17, with the exception of a small loss from 15 to 16 similar to the loss experienced by the boys at that age. At 17 the time required for girls was 31.5 hundredths of a second. Boys are better in this test than girls. The average of all the boys of all ages is 39.8 while that of the girls is 41. Not quite so much difference is seen here, however, as in the simple reaction-time where the average for boys was 20.2 while that of the girls was 22.3.

Columns *A*, *B* and *C* of table X show the length of time for bright, average and dull children respectively. In these results not quite so much difference is noticeable, which is perhaps due to the fact that they contain a somewhat smaller element of reaction-time, to which brightness and quickness are such close correlates. In this test, when all ages are considered, 40.1 is obtained as a result for bright

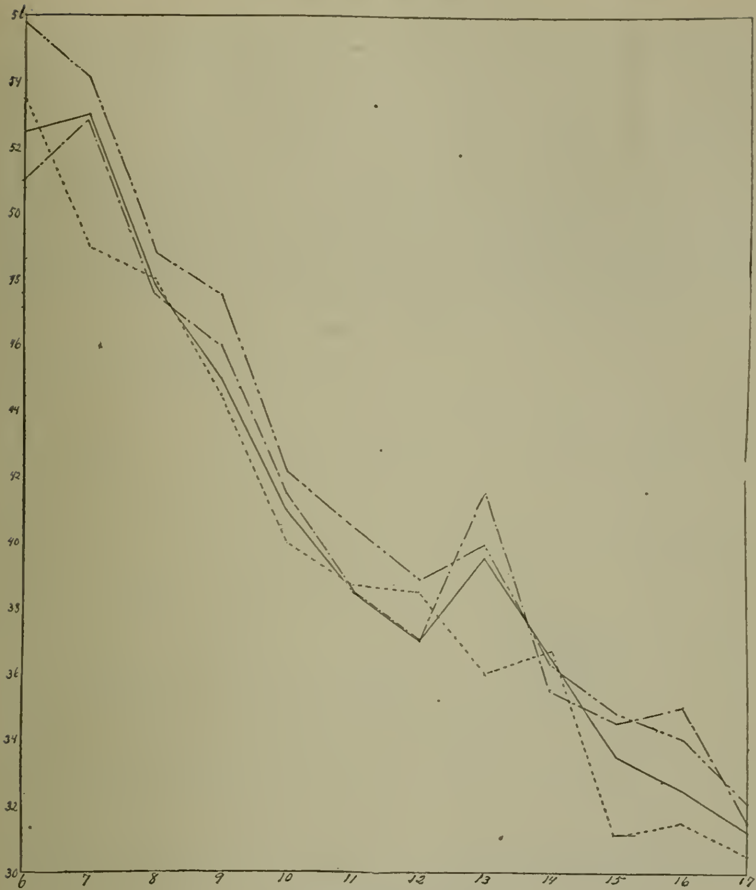

FIG. 25. CHART XX.

——————— Boys and girls.
.... Boys.
-—-—-—- Girls.
—--—--— Arithmetical averages—boys and girls.

children ; for average children, 40.2, and for the dull ones an average of 42.0 hundredths of a second. It is very evident from these figures that in this test the rule also applies that the brighter the child the more quickly he is able to react with discrimination and choice.

The average of the mean variations for separate children, shown by the dotted line, chart XXI, decreases gradually with age except from 12 to 14 where there is a marked increase, showing undoubtedly the effects of puberty during that period.

So far as the comparative mean variations for boys and girls are concerned, but little can be said, since the curves cross and re-cross

so frequently as to be of little value for comparisons between them and those in the main curve for development. The same general agreement exists between the main curve and the curves for variation in that there is increase in both with advance in years.

As explained under apparatus and methods the child was asked to react when blue appeared in the stimulating apparatus and not to react when red appeared. It is difficult for anybody to keep from reacting to the red, for the reaction becomes somewhat automatic and, if the attention flags, not infrequently reaction to red follows.

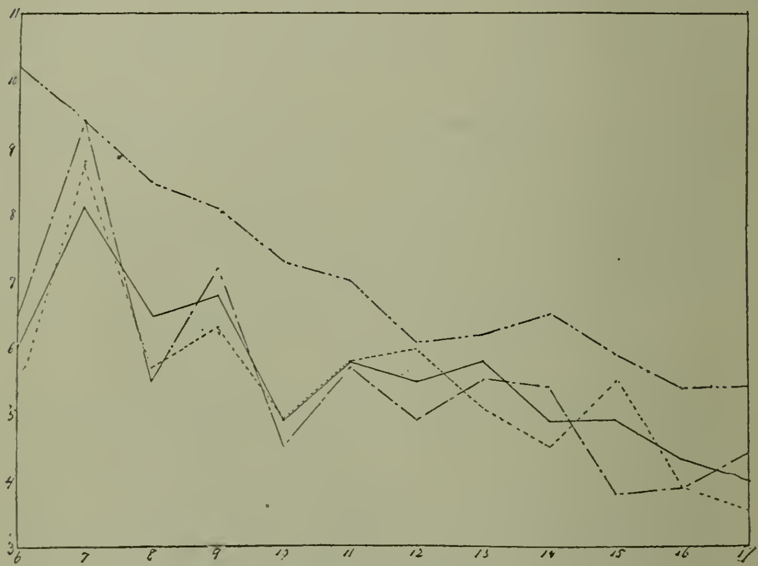

FIG. 26. CHART XXI.
——————— Statistical mean variation.
.... Statistical mean variation for boys.
-·—·—·—· Statistical mean variation for girls.
—··—··— Average individual mean variation.

Column E of table XI records the average number of reactions each child made to red out of the twenty trials given him. In order to be sure that the child discriminated between blue and red, as many reds had to be disclosed as blues, being irregularly mixed so that the numbers in column E would represent the number of mistakes made in 20 instead of 10, the latter of which would only represent the reactions to blue. Twenty of the 1192 children tested made the error of reacting to the red four times during the series of trials. Out of 1192 tested 420 made no errors at all by reacting to the red.

TEST (11): *Time-memory.*

As in the preceding tests, also in this one, ten trials were given each child and after taking the median value of these, the average of the mean variations for the separate results or data was calculated. Besides taking the median, the arithmetical average was taken for the sake of comparison. The arithmetical averages are to be found recorded in column *Ea* and the median values in column *Ep* of table XI. The median values for boys and girls are placed in columns *B* and *G* respectively. The averages of the mean variations for separate children are in the column headed *mv; MV* represents the mean variation for the total result, while *MV′* and *MV″* give the same for boys and girls respectively.

Columns *Ea* and *Ep* are used in constructing the two curves of chart XXV for the sake of comparing the results according to the two methods. Ages are at the bottom on the line of abscissas. The figures to the left indicate in hundredths of a second the amount of error made in making the second sound the same length as the first, which was two seconds long. The second sound was always made too short and the numbers thus indicate the amount of shortage in hundredths of a second.

The length of time, by using the arithmetical averages, is less than the median values from 6 to 10 years of age. Here, the curves, representing the result in graphic form, cross. From 10 to 15 the error is greater by arithmetical averages than by median values; from 15 to 16 the error is less; at 17 it is worse again. In calculating the results of this test the use of the method of median values is of still more importance than in the preceding tests, owing to the fact that the variations are sometimes extremely large. Frequently results were obtained in which all but one or two fell short of the correct time by 40 hundredths of a second or more while these two exceptions, owing to some disturbance or distraction either external or internal, were 20 or 25 hundredths too long. Accuracy with small mean variation demands a perfectly even flow of consciousness. During the first sound of two seconds the child simply sits and listens with no responsibility as to how long the sound shall go; all he has to do is to judge its length. When it becomes his turn to make the sound his responsibility and continuous wondering whether the sound is yet long enough make the time seem longer than it really is and consequently he stops it too soon.

The effect of suggestion in such a test is peculiar. For my tests the separate records of the child were, of course, kept secret until all

TABLE XI.

Time-memory.

Age.	*Ea*	*Ep*	*mv*	*MV*	*B*	*MV'*	*G*	*MV''*	*A*	*B*	*C*	*N*	*NB*	*NG*
6	56.7	62.0	24.6	23.4	56.5	25.1	67.0	23.2	55.5	68.5	69.5	94	52	42
7	59.6	66.5	27.9	20.2	63.5	20.6	68.5	20.4	66.8	66.5	67.3	96	49	47
8	52.7	54.3	23.6	22.8	48.5	22.3	57.0	22.3	48.5	59.3	69.0	96	49	47
9	56.2	60.0	23.0	23.5	47.5	22.4	73.5	19.3	50.0	65.0	46.5	97	49	48
10	48 9	48.5	20.2	18.1	48.5	21.8	46.5	15.8	40.3	49.5	69.5	95	49	46
11	44.2	41.0	20.8	18.2	40.5	16.6	41.0	20.2	45.0	42.0	48.0	101	50	51
12	41.6	36.8	17.6	21.3	35.8	21.8	37.5	18.7	28.5	49.3	44.5	106	56	50
13	36.3	33.0	17.9	21.4	24.5	22.0	36.0	19.5	36.0	25.8	33.5	110	51	59
14	35.9	30.0	18.7	16.1	31.5	14.2	31.0	17.8	28.3	28.8	42.5	103	49	54
15	37.6	38.0	18.0	19.4	34.5	15.3	39.0	21.9	22.5	36.0	45.0	102	51	51
16	41.6	44.0	16 6	16.7	38.0	16.5	49.0	14.0	43.3	39.3	47.0	87	48	39
17	39.9	35.5	13.8	15.8	34.0	13.8	40.0	17.8	33.5	38.8	31.3	91	47	44

Ea, hundredths of a second short in memory of two seconds—arithmetical averages.
Ep, hundredths of a second short in memory of two seconds—median values.
mv, average individual mean variation.
MV, statistical mean variation.
B, hundredths of a second short for boys.
MV', statistical mean variation for boys.
G, hundredths of a second short for girls.
MV'', statistical mean variation for girls.
A, hundredths of a second short for bright children.
B, hundredths of a second short for average children.
C, hundredths of a second short for dull children.
N, number of children.
NB, number of boys.
NG, number of girls.

had been taken. After taking the series of ten trials I again tried a number of individuals telling them each time the amount of error made. They soon learned to correct their error somewhat and not infrequently made the sound too long instead of too short.

A few of the younger children made the second sound not quite half as long as it should have been, making an error of more than 100 hundredths of a second. Only 38 out of the 1192 tested in all, judged the sound longer than it really was. It is interesting to note also that 19 out of these fell in the two ages 12 and 13, the former having 9 and the latter 10, none of the other ages having more than three each.

In time-memory the average child is worse at 7 than at 6, worse at 9 than at 8 and worse at 16 than at 14. The irregularity at the period of puberty falls later in the curves of this test than in the others. After the second decrease in ability from 8 to 9 there is a rapid increase in ability till 14. Thereafter it is reversed and there is more rapid loss than there was previous gain, leaving the child con-

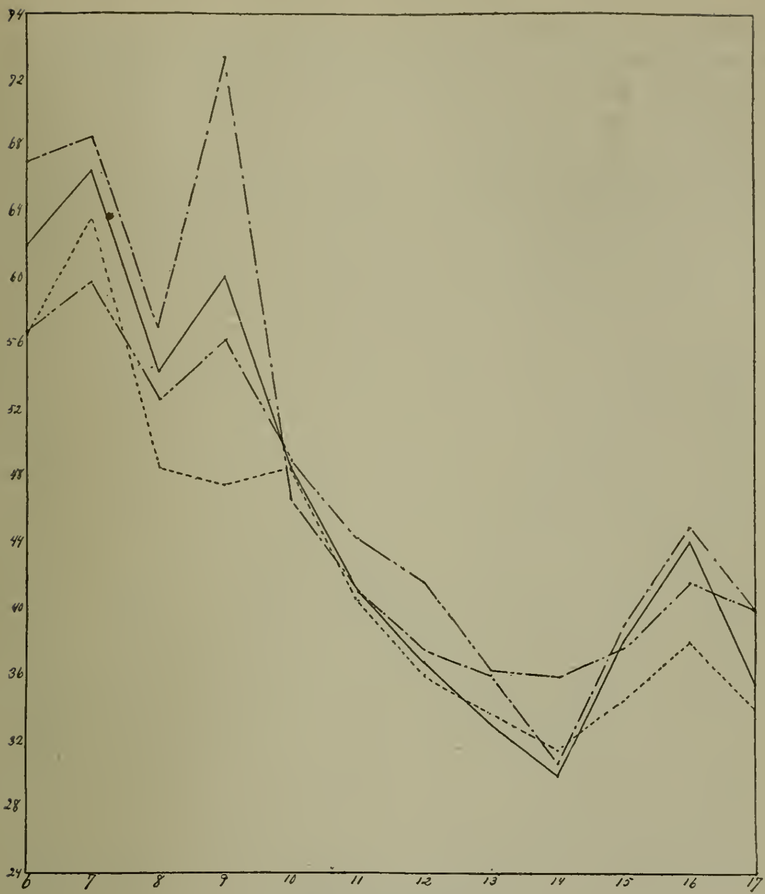

FIG. 27. CHART XXII.

——————— Boys and girls.
.... Boys.
- — - — - — - Girls.
— - - — - - — Arithmetical averages.

siderably worse at 16 than he was at 12. From 16 to 17 there is another rapid increase in accuracy. At 6 the child is 62 hundredths in error; at 7 the error is 66.5 ; at 8 it is 54.3. After rising again at 9 to 60 there comes the rapid increase in accuracy till the best point is reached at 14 where there is only an error of 30 hundredths of a second.

The results, when divided into boys and girls, show two very distinct differences. Both boys and girls lose from 6 to 7. With the exception of a very slight change in the average increase from 8 to

10, boys grow better till the age 13. Girls from 8 to 9 undergo an enormous loss, making the sound at the latter age 73.5 hundredths of a second too short. Thus, at this age they are far worse than at any other period between 6 and 17. From 9 to 10 an exceedingly rapid gain is made and thereafter there is a continuous gain till 14, instead of stopping at 13 as the boys did. Boys and girls both lose from this point till 16 where they again begin gaining. At 13 is the best time for boys, their error being only 24.5 hundredths too short; girls reach their best point at 14 with an error of 31 hundredths of a

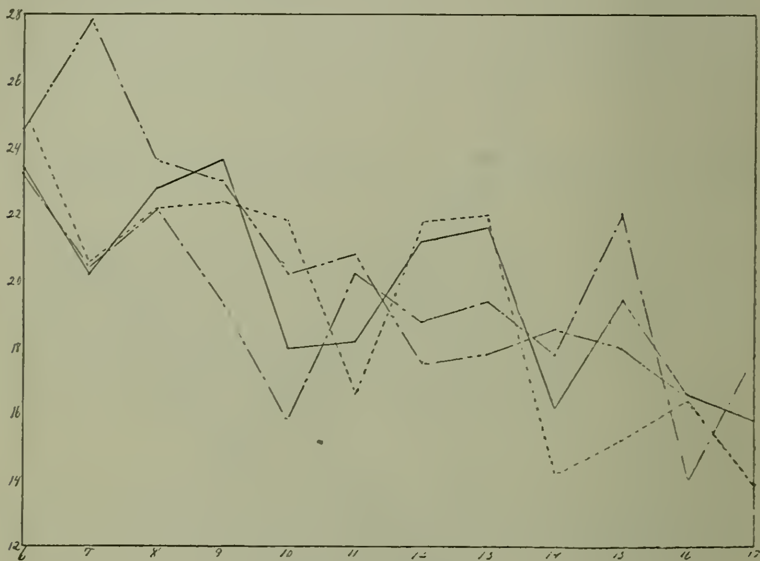

FIG. 28. CHART XXIII.
——— Statistical mean variation.
------- Statistical mean variation for boys.
-·—·—·— Statistical mean variation for girls.
——---— Average individual mean variation.

second. In this test also, boys greatly excel the girls, as can be seen by comparing columns *B* and *G* of table XI and also the curves of chart XXII.

In recalculating the results for comparison between bright, average and dull children the amounts of error made by them are 41.5, 47.4 and 51.1 respectively. The separate averages for each age and grade of children are recorded in columns *A*, *B* and *C* of table XI.

In general, ability increases with advance in years. This same is true of the curves of chart XXIII, representing the mean variations for time-memory. There is one thing here which is important as

corroborating the conclusion referred to under reaction-time, viz: that where a special change in relative growth between two ages occurs for the better, the mean variation changes for the worse. In chart XXIV the mean variation for both boys and girls decreases from 6 to 7, whereas the general ability for time-memory, chart XXII, increases for the corresponding period. The same can be noticed also in ages 8 to 10 and 14 to 16 where the most marked changes occurred.

General comparison of sex.

In order to get a general estimate of the mental differences of sex from the results obtained, the final averages of each age in muscle-sense, sensitiveness to color-differences, force of suggestion, reaction-time, reaction with discrimination and choice, and time-memory were thrown together into a general average for the respective ages. These general averages are recorded in table XII. The same are presented in the curves of chart XXIV. The ages are indicated at the bottom on the line of abscissas. The figures at the left can be given no definite value except to serve as relative points by which to judge the curves. The higher the figure the worse the record, just as was the case in the tests mentioned above. Voluntary motor ability and fatigue were not included in this general average both on account of their having a somewhat different nature from the rest, and, also one was the reverse of the other in that, in fatigue the lower the figure the better the record while the reverse was true for voluntary motor-ability. Could these two have been included, however, the result would have been all the more in favor of boys for it will be remembered that when these two were reduced to per cent. and considered together the balance was greatly in favor of boys. The boys became tired sooner but they also tapped much faster.

The only superiority shown for the girls is a small margin of advantage in discrimination for color-differences, referred to on page 48; the boys have the advantage very decidedly until between 7 and 8 but thereafter girls are better.

The general trend of the curves of chart XXIV is very interesting. Age 11 seems to be a neutral point where boys and girls are of about the same ability. From this age, the curves, on the whole, diverge in opposite directions more and more to the ages 6 and 17.

The general law of increase in ability is shown also very plainly by general averages taken in this way and represented in graphic form, as found in chart XXIV.

TABLE XII.

General comparison of sex.

Age.	*B+G*	*B*	*G*
6	35.1	34.1	36.4
7	36.4	34.3	36.8
8	32.9	31.6	33.3
9	33.0	30.8	35.5
10	28.4	27.7	28.5
11	25.8	25.7	25.8
12	24.5	24.1	25.0
13	23.5	21.4	24.6
14	21.9	21.5	22.1
15	22.8	21.3	24.0
16	23.2	21.4	25.4
17	20.1	19.2	22.0

The figures of the table are relative numbers deduced from the tests.

B+G, boys and girls.
B, boys.
G, girls.

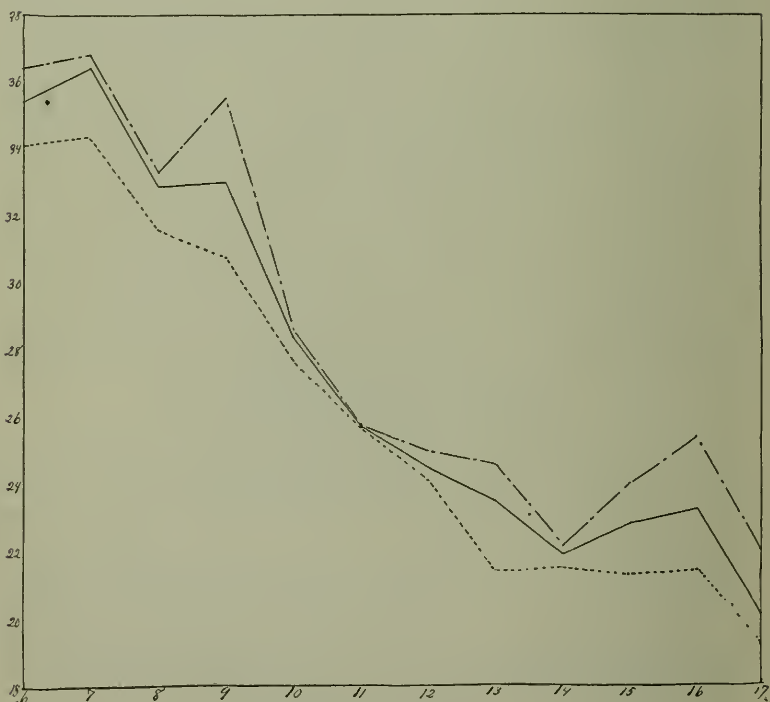

FIG. 29. CHART XXIV.

——— Boys and girls.
.... Boys.
—-—-— Girls.

INTER-RELATION OF THE RESULTS OF THE DIFFERENT TESTS.

The tests described in the preceding pages may be divided into two kinds, mental and physical. The close correlation of mind and body would naturally lead one to expect the closest sympathy between the two.

First, let us give a glance at the relations existing among the three physical curves, viz: weight, height and lung-capacity. Weight and height conform to almost exactly the same rules. In both, very slight differences exist between boys and girls until the period arrives at which boys begin to grow as men and girls as women. Boys are slightly heavier and taller than girls till between 11 and 12. Then the order is reversed. In both height and weight girls excel until between 14 and 15 where boys become heavier and taller and remain so the balance of life. The most rapid growth of girls ceases at 13 while at 14 the rapid growth for boys is just beginning. This difference between the growth of the sexes is all the more forcible when the curves for height and weight, charts XI and XIII, are compared with the one for lung-capacity, chart XV. After 12 the girls gain but very little in lung-capacity while the boys do not begin their real growth until 14. That the turning-point in life comes later for boys than for girls is verified not only by all the so-called physical curves but also by those which make the mental aspects more prominent. In all three physical curves there is the same general correspondence between the main curve* for each test and its accompanying curve for mean variations. In all, the variation increases with advance in years and when separate periods of each curve are considered the rate of increase or decrease in mean variations changes wherever there is a marked change in the rate of growth. The mean variations for boys and girls in the physical curves are largest during the years from 12 to 15.

In considering the so-called mental curves there is increase in ability with advance in years with the exception of the test on the force of suggestion. In this ability decreases till 9 in the sense that the suggestion is allowed to have more force; but, thereafter there is a gradual improvement as the suggestion decreases in strength.

At first sight the curve for mean variations of sensitiveness to color-differences, chart IV, seems to offer a marked exception to the general rule, that in all mental tests the mean variations decrease with advance in years. This is not the case. The mean variation rises from 6 to 9 because during those ages a large per cent. of the data were tens, which indicates that all ten colors seemed exactly

alike to those children. Consequently, the larger the per cent. of those picking out all ten as alike, the smaller the mean variation becomes for the data at hand ; but, could the 57 per cent. of non-discriminations at 6 be turned into the actual data which they could represent, the mean variation would be a great deal larger at 6 than what it is, viz: 1.8. Fifty-seven cases of non-discrimination necessitates 57 out of 100 of my data being the same: 10. This of course throws the mean variation very low.

In voluntary motor ability, chart VII, and fatigue, chart IX, more of the physical is involved, so that the relation between them and the two curves for weight and height is very marked. Voluntary motor ability from 12 to 14 is very probably affected by the very rapid growth for the corresponding period shown in charts XI and XIII. Separate relations exist for boys and girls between the curves for voluntary motor ability, fatigue and the three physical curves. The change in rapidity of growth for girls comes between 12 and 13 in weight, height and chest-capacity. The same change occurs at the same period in both voluntary motor ability and fatigue. For boys the marked change in rapidity of growth comes later than for girls, the former always changing radically at 14 as can be seen in charts XI, XIII and XV. In fatigue also the loss at puberty occurs later for boys than for girls, being at 14 instead of at 13 as it was for the girls. In voluntary motor ability the loss commences for both at the same place, viz: 12, but girls suffer loss for two successive years while boys lose only for one.

In voluntary motor ability the mean variation changes but slightly for different ages but in fatigue there is a more noticeable decrease in variation with advance in years.

The tests which are more strictly mental may be divided into two sections ; the first composed of tests (1), (2) and (3) and the second of tests (9), (10) and (11).

In muscle-sense, test (1), the effect of puberty on the rusults is very marked as is shown by chart I, but in discrimination for color-differences almost no divergences whatever can be noticed at that period. In the curve for force of suggestion, boys and girls alike suffer a loss in ability from 14 to 15. The curves of mean variations in these three tests all show marked divergences from the general trend during the years from 12 to 15. By throwing the muscle-sense and force of suggestion into relation, a purely mental element is brought out in the latter. Had the weight been subjected to the muscle-sense alone the discriminative ability by that sense would

always have said they were of equal weight, for, by test (1) and chart I, it is shown that discrimination for weight increases gradually with age, but by considering the element of sight we get a measure of our error in judgment in test (3), chart V.

In tests (9), (10) and (11), where more quickness and accuracy of action is involved, the effects of puberty show themselves far more plainly than in any of those hitherto considered. Thus, it might be concluded that puberty has a greater effect on the mental than upon the physical aspects of man's nature. The difference is far more noticeable in girls than in boys as can be readily seen by referring to charts XVII, XX and XXII. This is specially noticeable in chart XXIV for the comparison of sex. The development of girls is far more seriously affected by periodic changes than that of boys as can be seen by reference to 7, 9, 14 and 16 of this chart. In discrimination time, chart XX, the loss in ability for girls comes before that of boys, the former beginning at 12 and the latter at 13. This order is reversed in time-memory, chart XXII, and comes later for both, boys beginning to lose in accurrcy at 13 and girls not until 14. This difference between discrimination-time and time-memory is brought out all the more forcibly by considering the respective curves for boys and girls combined, charts XX and XXII. In the former the loss is between 12 and 13 ; in the latter it occurs between 14 and 16.

The mean variation for the separate ages in tests (9), (10) and (11) furnish but little suggestion as to the effect of puberty. The averages of the individual mean variations for separate children show marked changes from 12 to 14 in discrimination and time-memory but nothing in simple reaction-time as is shown by the dash double dot lines of charts XVIII and XXVII.

In all the curves involving a time-element there is marked loss in ability from 6 to 7 as can be seen in charts XVII, XX, XXII and IX. In reaction-time and reaction with discrimination and choice, charts XVII and XX, the loss is only experienced by girls, but in the other two, time-memory and fatigue both alike suffer some set-back from 6 to 7. The same thing is noticeable to a less degree in muscle-sense and color discrimination, charts I and III, the boys alone suffering loss in color-sensitiveness. The general fact, however, is brought out very forcibly in comparative curves of sex, chart XXIV.

An interesting fact is brought out by throwing graded reaction and the same with discrimination and choice into relation with each other.

Between 11 and 12, just before puberty, in both curves the bright and dull children act with about the same rapidity, though both before and after that age the dull ones are much slower than the bright ones. It is evident from a glance at the two charts XIX and XXII that the child is judged dull at 13 because he is unable to act as quickly. Considering these charts in conjunction with the comparative curves of sex, chart XXIV, the conclusion is suggested that all children on an average are of about equal ability at age 11.

Relation of individual tests to general mental ability.

The data for the tests on weight, height, lung-capacity, discrimination-time, reaction-time and time-memory were recalculated to get the relation between the separate tests and general mental ability as estimated by the teacher. The curves for height and reaction time were inserted in part II under tests (7) and (10) respectively. The curves for height, worked out in relation to " stand " are found in chart XV and those for reaction time in chart XIX. The remainder of the curves belonging to this class have been given a separate section on account of their largely negative value. It will be remembered that the curves for height were of negative value in that no marked relation could be traced between them and the mental ability of the pupils upon whose measurement they were plotted. The same is true of weight ; also of chest capacity with the exception of the years 13 and 14 where the dull pupils have a much smaller lung-capacity. The curves for reaction-time gave the most positive results showing that the brighter the child the more quickly he is able to act. In discrimination, the same relation is noticeable but to a less degree.

The bright children and those of average ability, as judged by the teachers, are about equal in the length of time required to discriminate but the dull ones require a somewhat longer time at all ages with the exception of 9, 11 and 13. The difference here between the classes is not so great owing probably to the fact that it involves a smaller element of reaction-time which shows the most marked difference. In turning to the last one, viz : time-memory, it may be said in general that the brighter the child the more accurate his sense of time. It is impossible to say that this is the case with all ages, because the curves, as can be seen, cross and re-cross so frequently as to be of but little value further than to indicate the relation of the three grades in a very general way.

COMPARISONS WITH THE RESULTS OF OTHER INVESTIGATORS.

Owing to the limited amount of investigation on the mental development of school-children but little can be said as to how my mental tests agree with other investigations. For the purely physical tests, weight, height and lung-capacity, more material presents itself. For purposes of comparison with the first two—weight and height—I have chosen the results of BOWDITCH[1] of Boston and of PECKHAM[2] of Milwaukee. The weights of the children in both these investigations as well as in mine were taken in ordinary clothes. My original calculations were made by the method of median values and the curves previously given are plotted upon this method. Since the results of both BOWDITCH and PECKHAM were obtained by arithmetical averages, for sake of comparison of the two methods, and also for comparison of localities, I recalculated my results by arithmetical

TABLE XIII.

Comparative weights of Boston, Milwaukee and New Haven school-children.

Age.	*B*	*M*	*NHa*	*NHp*	*B*	*M*	*NHa*	*NHp*
	Boys.				*Girls.*			
6	45.17	44.81	47.86	46.8	43.23	43.12	45.92	44.3
7	49.07	49.10	52.08	51.2	47.46	46 97	49.65	50.4
8	53.92	53.81	56.49	52.5	52.04	50.87	53.32	53.0
9	59.23	59.46	61.69	60.0	57.07	56.44	59.27	58.8
10	65.30	65.35	68 31	68.4	62.35	62.45	66.09	62.7
11	70.18	70.92	75.57	70.8	68.84	68.84	72.35	70.0
12	76.92	76.08	83.25	82.3	78.31	77.82	86.09	84.5
13	84.84	84.89	90.95	88.0	88.65	87.96	92.97	92.0
14	94.91	95.76	97.45	91.7	98.43	97.64	97.15	98.0
15	107.10	109.05	109.95	110.0	106.08	105.87	103.06	104.0
16	121.01	122.06	126.03	127.0	112.03	110.58	111.41	113.0
17	127.49	130.35	126.70	130.0	115.53	113.32	118.45	113.7

B, weight of Boston school-children in pounds.

M, weight of Milwaukee school-children in pounds.

NHa, weight of New Haven school-children in pounds—arithmetical averages.

NHp, weight of New Haven school-children in pounds—median values.

[1] BOWDITCH, *Growth of children*, VIII. Ann. Rept. State Board Health Mass., 307, Boston 1877.

BOWDITCH, *Growth of children*, X. Ann. Rept. State Board Health Mass., 35, Boston 1879.

BOWDITCH, *Growth of children studied by* GALTON'S *method of percentile grades*, XXII. Ann. Rept. State Board Health Mass., 479, Boston 1891.

[2] PECKHAM, *Growth of children*, VI. Ann. Rept. State Board Health Wisconsin, 28, Milwaukee 1881.

averages. Notwithstanding the protest of Peckham against the median, in all mental tests, at least for getting a result for each child, it is far preferable to the arithmetical averages.

Peckham found that Milwaukee children were heavier and taller than Boston children. New Haven children are shown by my results to be still heavier than either of the other two. The comparative weights in pounds are to be found in table XIII and also

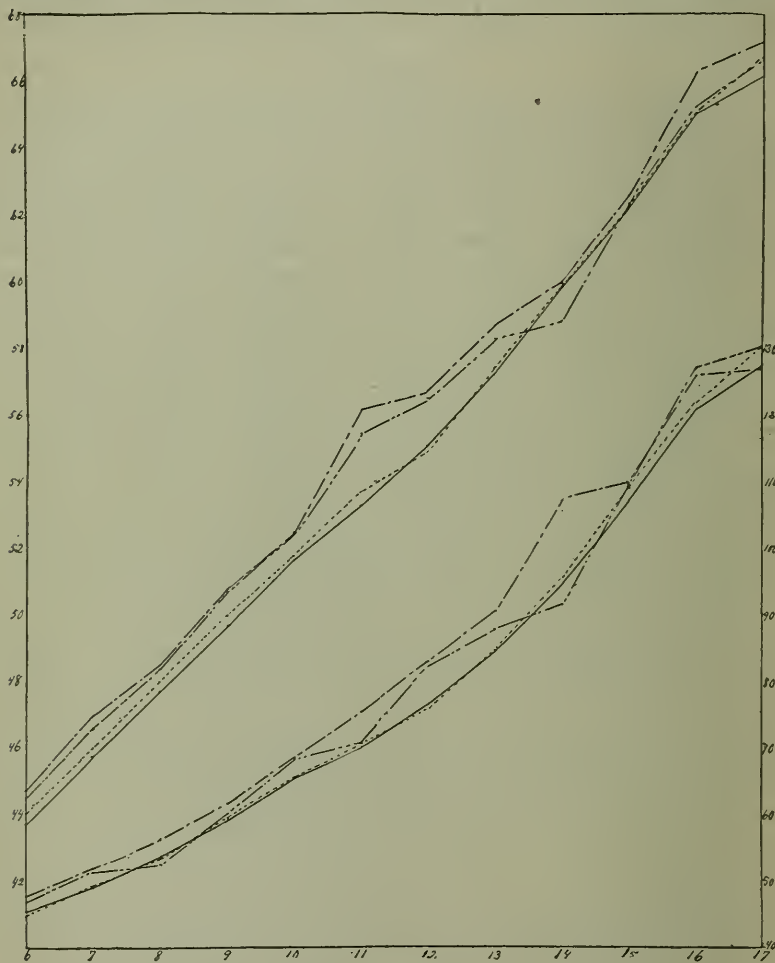

Fig. 30. Chart XXV.

Lower curves for weight.
Upper curves for height.
———— Boston boys.
.... Milwaukee boys.
-.-.-.- New Haven boys.
—..—..— New Haven boys—median values.

in graphic form in charts XXV and XXVI. As no relation between the arithmetical averages and median values was considered above, I have inserted here the curve of median values, as well as that of arithmetical averages. Ages are at the bottom; the figures to the left indicate height in inches; those to the right indicate weight in pounds.

The height of the children in my tests was taken with shoes; those of Bowditch and Peckham were taken without shoes. After subtracting the height of an average heel of a shoe, which is about 1 inch, the New Haven children are still taller than those of Boston and Milwaukee. The results for height are recorded in table XIV and charts XXV and XXVI.

Table XIV.

Comparative heights of Boston, Milwaukee and New Haven school-children.

Age.	*B*	*M*	*NHa*	*NHp*	*B*	*M*	*NHa*	*NHp*
	Boys.				*Girls.*			
6	43.75	44.08	45.2	45.0	43.35	43.78	45.0	44.9
7	45.74	46.09	47.4	47.1	45.52	45.93	46.9	46.9
8	47 76	48.05	49.0	48.9	47.58	47.59	48.8	48.4
9	49.69	50.00	51.3	51.2	49.37	49.81	51.1	50.8
10	51.68	51.85	53.0	53.0	51.34	51.89	53.2	52.8
11	53.33	53.76	56.7	55.9	53.42	53.80	54.1	54.6
12	55.11	54.98	57.2	57.0	55.88	56.47	58.3	57.9
13	57.21	57.47	59.2	58.8	58.16	58 68	59.6	60.4
14	59.98	59.89	60.4	59.3	69.94	60.50	60.7	61.4
15	62.30	62.34	63.1	62.8	61.10	61.59	62.2	62.5
16	65.00	65.07	66.7	65.7	61.59	62.16	62.8	62.5
17	66.16	66.60	67.1	66.1	61.92	62.91	63.8	63.6

B, height of Boston school-children in inches.

M, height of Milwaukee school-children in inches.

NHa, height of New Haven school-children in inches—arithmetical averages.

NHp, height of New Haven school-children in inches—median values.

This difference in weight and height is very likely due to the smaller proportion of foreigners included in my results. American-born children are taller and heavier than foreign-born children.[1]

So far as the relation of growth of different ages is concerned, the same general laws appear in my results as those obtained by Bowditch and Peckham.

It is interesting to notice the effect of the private-school gymnasium upon the development of lung-capacity. Table XV and chart XXVII

[1] Bowditch, *Growth of children*, VIII. Ann. Rept. State Board Health Mass., 307, Boston 1877.

represent a comparison of my results taken in public schools with the results of Anderson, taken in private-schools near New York. His pupils underwent a daily training in the gymnasium. The boys in his results start at 6 better than those of the same age in my results. They also develop more rapidly. No loss at puberty is traceable, however, in his results similar to that shown in mine. His results extend only to 15 years of age. At this age private-school boys had a lung-capacity of 205 cubic inches; public-school boys by my results only had a capacity of 170.3 cubic inches, and only 202.5 cubic inches at 17, which is 2.5 cubic inches less than private-school boys had at 15. In girls the difference is even more noticeable, for at 6 the girls

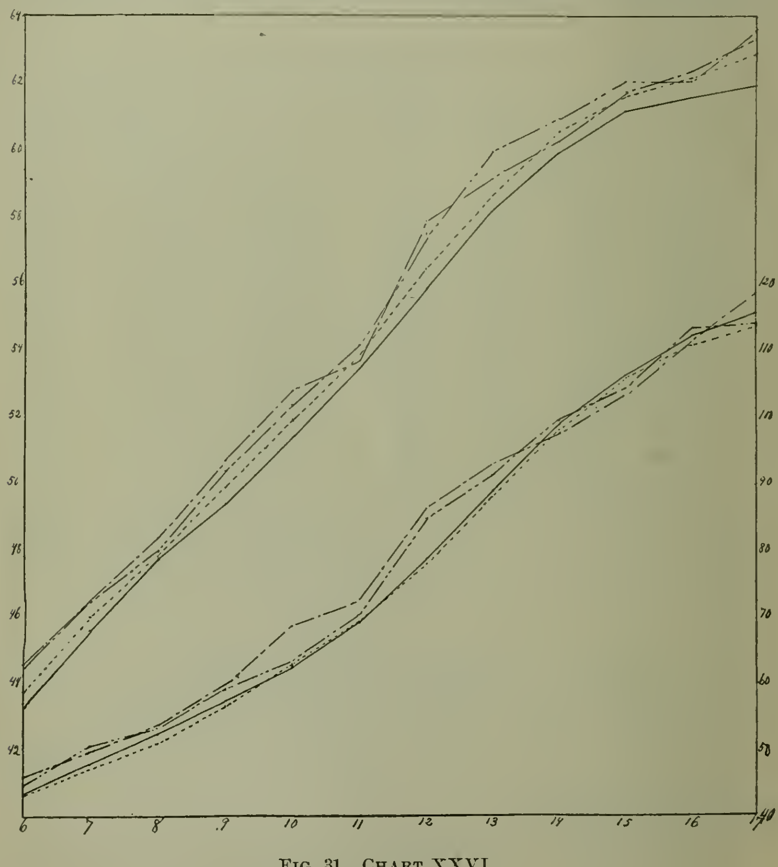

Fig. 31. Chart XXVI.

Lower curves for weight.
Upper curves for height.
———— Boston girls.
.... Milwaukee girls.
-—-—-—- New Haven girls.
—------— New Haven girls—median values.

TABLE XV.

Comparative lung-capacities for public- and private-school children.

Age.	*P*	*NHa*	*NHp*	*P*	*NHa*	*NHp*
	Boys.			*Girls.*		
6	64	57.1	56.0	35	49.2	50.0
7	80	65.6	66.0	40	58.9	54.0
8	88	71.8	73.0	48	64.2	66.0
9	106	82.7	83.0	65	73.4	72.5
10	124	92.8	91.5	80	80.3	82.0
11	144	106.3	104.0	106	83.6	83.0
12	150	117.2	113.5	125	104.6	104.0
13	168	124.4	120.0	136	108.1	105.0
14	188	120.4	125.0	150	107.8	105.0
15	205	170.3	161.0	155	116.3	116.0
16	---	189.3	187.0	---	119.9	115.0
17	---	202.5	204.0	---	124.0	118.5

P, lung-capacity of New York private-school children.

NHa, lung-capacity of New Haven public-school children—arithmetical averages.

NHp, lung-capacity of New Haven public-school children—median values.

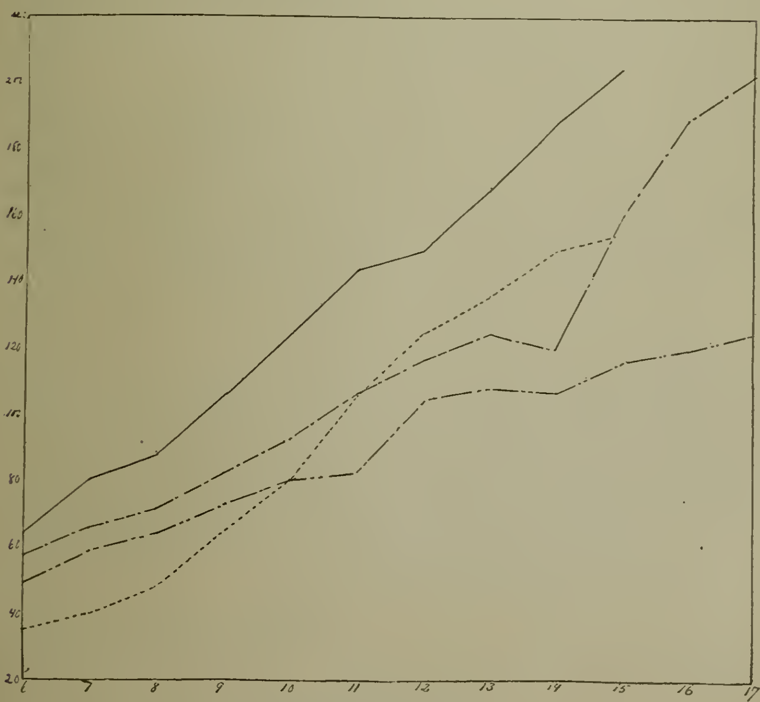

FIG. 32. CHART XXVII.

Private school boys.
Private school girls.
Public school boys.
Public school girls.

at private schools have a smaller lung-capacity but develop so rapidly as at 10 years of age to equal those of public schools. Thereafter the rapid development continues so that at 17 they have attained 155 cubic inches while my results only show 116.3 cubic inches for public-school girls; even at 17, public-school girls have only 124 cubic inches capacity, which is 31 cubic inches less than that of private-school girls at 15. The curves of chart XXVII show an enormous difference between the two classes of children. There is added, however, a note to ANDERSON's table: "It cannot be said of them that they indicate just what the averages should be." The results of this table, however, were averages of about 600 children of each age.

In chart XXVII the ages are at the bottom and the figures at the left indicate the lung-capacity in cubic inches.

REMARKS ON DR. GILBERT'S ARTICLE,

BY

E. W. SCRIPTURE.

In preparing for publication Dr. GILBERT's *Researches on the mental and physical development of school-children*, several matters seemed to me to require a further statement.

Suggestion-test. The large and the small standard are of the same weight, namely, $w=55^g$. The diameter of the small standard was $d=2.2^{cm}$, and of the large standard $D=8.2^m$. The weights of the 14 blocks were successively

$$p_1=15^g;\ p_2=20^g;\ .\ \ .\ \ .;\ p_{14}=80^g,$$

the difference between any two successive blocks being $p_i-p_j=5^g$. The diameter was $\delta=3.5^g$ for all. The result of the child's two judgments was that $w=p_k$ for the small standard and $w=p_l$ for the large standard. The amounts of difference $v_k=p_k-w$ and $v_l=p_l-w$ can be taken as measures of the effect of the different sizes of the standard and the blocks of the 14-series. As all blocks were cylinders of the same length, we can, if we neglect effects of contrast of length to diameter, express the difference in size by the areas of the ends. Thus the visual suggestions can be indicated by $\frac{1}{4}(\pi\delta^2-\pi d^2)=u_k$ and $\frac{1}{4}(\pi D^2-\pi\delta^2)=u_l$ respectively. As the differences v_k and v_l disappear when the blocks are lifted without being seen, we can put

$$v_k=f_k(u_k) \text{ and } v_l=f_l(u_l).$$

Full expressions for f_k and f_l for a given individual would give the law of suggestion for the given case. The determination of this law was not the object of the investigation, in which u_k and u_l were taken constant. For the sake of brevity the difference between the large block and the small block was taken as the amount of suggestion; thus the constant suggestion was taken as

$$S=u_k+u_l=\tfrac{1}{4}(\pi D^2-\pi d^2).$$

The result of this suggestion is

$$H=v_k+v_l.$$

If by A we indicate the age then

$$H=f(S, A),$$

and by taking $S=$ constant, we obtain

$$H=f(A)$$

or the force of suggestion as a dependent on age.

Method of computation. The method of computing the results can be thus indicated. Let the results be

$$\begin{array}{llll} a_1, & a_2, & \ldots, & a_n \\ b_1, & b_2, & \ldots, & b_n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ l_1, & l_2, & \ldots, & l_n \end{array}$$

where the letters $a, b, \ldots, l$ refer to different children and the indices $1, 2, \ldots, n$ refer to experiments on the same child.

If we express by $R_x=f_i(x)$ the fact that R is determined from the i powers of x, then for the median we have the general expression

$$M_x=f_0(x_1, x_2, \ldots, x_n)$$

and for the arithmetic mean

$$A_x=f_1(x_1, x_2, \ldots, x_n).$$

For the reasons indicated on p. 23 the mean variations are determined as f_1.

We have thus for the individual children the medians

$$\begin{array}{l} c_a=f_0(a_1, a_2, \ldots, a_n) \\ c_b=f_0(b_1, b_2, \ldots, b_n) \\ \cdot \\ \cdot \\ \cdot \\ c_l=f_0(l_1, l_2, \ldots, l_n) \end{array}$$

The mean variations for the individual children will be

$$\begin{array}{l} D_a=f_1([a_1-c_a], [a_2-c_a], \ldots, [a_n-c_a]) \\ D_b=f_1([b_1-c_b], [b_2-c_b], \ldots, [b_n-c_b]) \\ \cdot \\ \cdot \\ \cdot \\ D_l=f_1([l_1-c_l], [l_2-c_l], \ldots, [l_n-c_l]) \end{array}$$

These mean variations can be regarded as psychological quantities expressing the accuracy of each child's judgment. The median accuracy of judgment for the particular age r will be

$$D_r = f_0(D_a, D_b, \ldots, D_l).$$

This can be called the mean personal insecurity.

These individual quantities are quite different from the statistical mean variations, which are taken in order to show the homogeneity of the children for any given age. The median for a given age r is found by

$$C_r = f_0(c_a, c_b, \ldots, c_l).$$

The mean variation of the children from the general mean will be

$$D_r = f_1([c_a - c_r], [c_b - c_r], \ldots, [c_l - c_r]).$$

Limited series. In tests (1), (2) and (3) it was discovered too late that the series of blocks and colors did not extend far enough to include extreme cases. In each table there is a column with the percentage of cases where the least perceptible difference or the force of suggestion went beyond the limits of the apparatus, and in the charts dotted curves are given for these cases. GILBERT states that the column of mean values does not give the quite correct result unless it be considered in relation to the column of percentages just mentioned. This statement would be true if the mean value used had been the arithmetic mean but is not true for GILBERT's results as the mean value used was the median. As explained on p. 32 the separate values influence the median only as being above or below it. The median child remains just the same whether an extreme child exceeds the recording power of the apparatus or not. The columns of means and their curves are thus completely correct in themselves. The columns of percentages of no-discrimination give really another representative value of no particular use in itself but quite important when compared with the medians as indicating the form of the frequency-curve and the even course of the curve of results according to age. In this respect it is as important as the mean variation.

The column of these percentages, although having no influence on the median, does have an influence on the mean variation. The mean variations are all too small by a quantity ζ following the law $\zeta = f(P)$. If the positive and negative mean variations had been calculated for the $(100-P)\%$ of the cases separately, a deduction on the assumption of (11), p. 12, would have rendered it possible to

calculate the actual mean variation for the whole number independent of the limitations of the apparatus. The gain, however, would have been incommensurate with the labor; the mean-variation-curves would resemble those actually given and would simply be steeper at the left.

The columns headed PB and PG give us the means of answering the question as to when the difference between the boys and girls recorded for any given age is to be considered as a true difference between the sexes or as merely the result of the finite number of cases considered.

The method is as follows. Let the percentage of the total n boys within the limits of the test be $p=(100-PB)\%$ and of those beyond be $q=PB\%$; likewise for n girls let the percentages be $p'=(100-PG)\%$, $q'=PG\%$. According to a well-known theorem[1] we can assert with a probability of $\Phi(\gamma)$ that the two classes are different, provided

$$p-p' > \pm\gamma\sqrt{\frac{2pq}{n}+\frac{2p'q'}{n'}},$$

where $\Phi(\gamma)$ is the function expressed on p. 20.

I have tested some specimen cases in this way and find that for these three tests it cannot be asserted in many cases with a practical certainty of $\Phi(\gamma)=0.999978$ that there is a real difference between the two sexes.

Another test is furnished by Bayes's theorem used on p. 38. This theorem can be applied to all the tables.

In a like manner the differences between successive ages can be tested.

---

[1] Illustrated in Lexis, Einleitung in die Theorie der Bevölkerungstatistik, 103. Strassburg 1875.

EXPERIMENTS ON THE HIGHEST AUDIBLE TONE,

BY

E. W. Scripture and Howard F. Smith.

The highest audible tone, or the upper limit of pitch, is that tone at the extremity of the series of tones arranged according to pitch beyond which any rise in frequency of vibration fails to produce a sensation of tone. The highest audible tone is defined psychophysically by the frequency of the physical vibrations corresponding to it. The term "vibration" is understood to mean one complete pendular oscillation including both phases. The highest audible tone has been differently determined by various observers: Sauveur, 6 400; Chladni, 8 192; Wollaston, 25 000; Savart, 24 000; Despretz, 36 864; Blake, 40 000 to 60 000.

The great discrepancy in the results is usually said to have been due to the imperfections of the apparatus employed. There is no need, however, of this assumption, as there is a source of variation quite sufficient to explain the discrepancy; it is unquestionable that these men worked with tones of different intensities.

Even in the very latest experiments the factor of intensity has been generally overlooked. Savart was the first to observe that the highest audible tone was different for different intensities. Rayleigh[1] took care to keep his tones of approximately the same intensity. Blake, who used a succession of steel bars of varying length, produced the tone by a pendulum-hammer swinging over a graduated scale, thus insuring a nearly uniform stroke and correspondingly uniform intensity of tone. These experiments, however, went no farther than to secure a constant intensity.

The highest audible tone requires in each case a measurement of intensity as well as of pitch in order to complete its determination. It thus becomes important to inquire how the pitch of this tone depends on the intensity, or, in other words, to determine what the highest audible tone is for each intensity.

Apparatus.

The first step in solving this problem was the selection of an apparatus giving tones of the required pitch, but so arranged that the

[1] Rayleigh, *Acoustical observations; Very high notes*, Phil. Mag., 1882 (5), XIII 344.

intensity could be kept constant at any desired point and could be readily and accurately varied. Tuning forks were not used because the range for the highest audible tone is so wide that a very large number of forks would be required, and also because they do not admit any accurate regulation of intensity. The rods of König with pendulum-hammer are in some respects better than the forks but are not very easy to manipulate in rapidly conducted tests, such as are necessary in order not to fatigue the observer ; moreover, the sound of the impact of the hammer cannot fail to be a disturbing element in the experiment. Neither the forks nor the rods give a sustained tone, but one of rapidly decreasing intensity.

The Galton-Koenig whistle was selected as the most reliable and readily manipulated instrument for the tests proposed. It gives a sustained tone as long as the blast of air continues ; the pitch of the tone is readily and accurately varied. The whistle consists of a brass tube, about 7^{cm} in length with a cap screwing over one end. To this cap there is attached an accurately fitting piston, which moves inside the tube as the cap is screwed upward or downward. The outside of the tube is marked with a longitudinal scale the unit of which is 1^{mm}, the length of the scale being 12^{mm}. The screw is so arranged that one complete turn of the cap carries the piston a distance of 1^{mm}. The upper rim of the cap is divided into ten divisions, each one representing a movement of the piston through 0.1^{mm}. These divisions being very large, the sub-divisions of 0.01^{mm} can be obtained by the eye without error. The maker's graduation was verified to 0.01^{mm}. The whistle is blown by a current of air forced into the bottom of the tube ; near the bottom there is a narrow slot extending across one-half of the circumference. The whistle is thus a closed labial pipe and the tone produced will be determined by

$$n = \frac{v}{4l},$$

n being the number of vibrations, v the velocity of sound and l the length of the pipe. The velocity of sound in dry air at 0° C. is generally given as 330.7^{m}; for the temperature of $t°$ it will be

$$330.7\sqrt{1 + 0.00367\ t.}$$

The average temperature of the room used can be taken as 20° C. ; as the temperature was not recorded we can assume $\pm$ 1° C. as the limit of fluctuation during a set of experiments. This gives $v = 342.525^{m}$ with a mean error of 5 per cent. (estimated). As the actual

mean variation for a set of results seldom exceeded 5 per cent. it may well be supposed that the constant of precision of the measurements was in this case determined by technical errors and not by psychological variations. In future experiments it will be necessary to maintain the room at an even temperature during each set of experiments; and to calculate the pitch with the appropriate value of v for each change, possibly also to be on guard against sudden barometric changes.

It may seem remarkable that we should have neglected to record the temperature of the room. We give the following as reasons: 1. we did not expect after the elimination of the error of air-pressure to find the psychological sources of error smaller than the technical ones; 2. the temperature of the air has not been regarded in previous experiments; 3. it is not the custom of psychological laboratories to pay attention to the psychological and instrumental errors due to changes in temperature.

Several means of blowing the whistle have been employed. GALTON[1] used a single rubber bulb. ZWAARDEMAKER[2] used a funnel with a rubber membrane stretched across the large opening, the smaller end being connected with the whistle by a rubber tube; the funnel was depressed through a constant distance. These methods are none of them strictly reliable. The pressure is intermittent and cannot be accurately regulated. The only possible method of obtaining a current of air of constant intensity, seems to be by use of a rotary-fan blower. This method has been previously described and tested.[3] The source of power used by us was an electric motor run by the city-current; there were no perceptible fluctuations in speed. The fan-wheel of the blower made from 13 000 to 15 000 revolutions a minute. The blast was carried by a rubber hose into a room in another part of the laboratory; thus all noise from the machinery was avoided. The hose led to a rubber tube, in which was a stop-cock. The rubber tube ended in a glass T-coupling with a rubber tube on each end of the cross-arm. One of these led to the whistle, the other to a water-manometer. The manometer scale was graduated to millimeters; the height of the column of water gave the

---

[1] GALTON, *Whistles for audibility of shrill notes*, Inquiries into Human Faculty, 38, New York 1883.

[2] ZWAARDEMAKER, *Der Umfang des Gehörs in den verschiedenen Lebensjahren*, Zt. f. Psych. u. Phys. Sinn., 1894 VII 10.

[3] SCRIPTURE, *A constant blast for acoustical purposes*, Am. Jour. Psych., 1892 IV 582.

pressure of the blast of air supplied to the whistle. The pressure was regulated by the stopcock. In this manner a constant blast of air could be maintained for any given time and the intensity could be varied at will. The fluctuations of pressure as indicated by the manometer did not exceed 1 per cent. Experiments were made with five different pressures, 50^{mm}, 100^{mm}, 150^{mm}, 200^{mm} and 250^{mm}. Within the limits of accuracy of 5 per cent. with which the experiments were conducted, the intensity of the vibratory movement could be considered as varying in direct proportion to the pressure of the blast.

In some of the earlier experiments we used a foot-blower as a bellows. Even the best bellows cannot be accurately regulated ; a foot-blower still less so. The following mean variations are from experiments on trustworthy observers. With the foot-blower we find, for example, mean variations of 10%, 14%, 5%, 9%, 8%, 13% ; with the rotary-fan blower, 5%, 9%, 2%, 7%, 4%, 2%. The records are not on the same observers in the two cases but the difference is sufficient to indicate quite a gain in accuracy by using the rotary blower.

The whistle was started near 0, i. e. above the upper limit of pitch, and the cap was gradually unscrewed until the observer detected a tone. The reading of the scale was observed at this point. The cap was unscrewed a little further to make sure that a good musical tone was heard; then it was screwed up again, stopping at the instant the observer lost the tone. The two readings were noted down in separate sets, D and A. This was repeated five times, making 10 records with the given intensity. Sets of records were made in succession with the five pressures, beginning at 50^{mm}. After a rest of about 3 minutes another series of 50 records was made but the pressures were used in the reverse order. This reversal of the order of the pressures eliminated the error of fatigue, if there was any.

It will be noticed that the method differed from that used by previous observers, being the method of regular variation.[1]

The execution of the experiments was in charge of H. F. Smith who is responsible for the care exercised. The setting up of the blower, shafting and belting was done by the laboratory mechanic, J. H. Hogan, who controlled the running of the machinery during the experiments. Rubber belting was used, but owing to its inferior flexibility it should in future be avoided for small pulleys like that

---

[1] SCRIPTURE, *On the method of regular variation*, Am. Jour. Psych., 1891 IV 577.

SCRIPTURE, *Ueber die Aenderungsempfindlichkeit*, Zt. Psych. Phys. Sinn., 1894 VI 472.

on the blower ; with the high speed used it has been ignited by the slipping. The whole arrangement of apparatus was supervised and inspected by E. W. Scripture.

Computation.

The median value was taken for each single set of five of the same kind. For the result descending or for the result ascending the two medians were weighted in the usual way inversely as the square root of the mean variations and were then averaged. For the final result the four medians were weighted and averaged.

As the use of the median is new, all the results were calculated also in the usual way to obtain the arithmetical mean. The components were weighted in the regular way.

The computation was done by E. W. Scripture with the aid of Crelle's, Rechentafeln. There were in all 120 sets of 5 results each. The median of each set was first recorded, then the error of each result from its median was written underneath it. The average of each set and the error of each result were likewise computed and were written beside the median and its errors. It was soon noticed that the average seldom differed much from the median ; this was used as a check on the computation of the average, the whole computation being repeated in case of much divergence. It was also noticed that, given the median and the average, the errors from the average could be calculated from the column of errors from the median ; for the latter half of the work this was done by the amanuensis while the computer calculated the errors directly. As a credential for the reliability of the computation, it can be said that only one mistake in the calculation of the errors was found in the results verified in this way. The mean errors of each of the 240 sets of five errors were computed according to the formula on pp. 18, 24, the division by 4.5 being performed from the multiplication table for 45. The square root of each mean error was taken from a table of square roots. The reciprocals for weights were written as decimals to two places.

Influence of intensity.

The results are given in table I. Although the results were calculated into vibrations, they were allowed to stand in the table as hundredths of a millimeter in order to avoid the impression of a false degree of accuracy which arises when 0.01^{mm} is turned into a whole number with zeros at the end. In the curve (fig. 33) the frequencies are given for the sake of comparison.

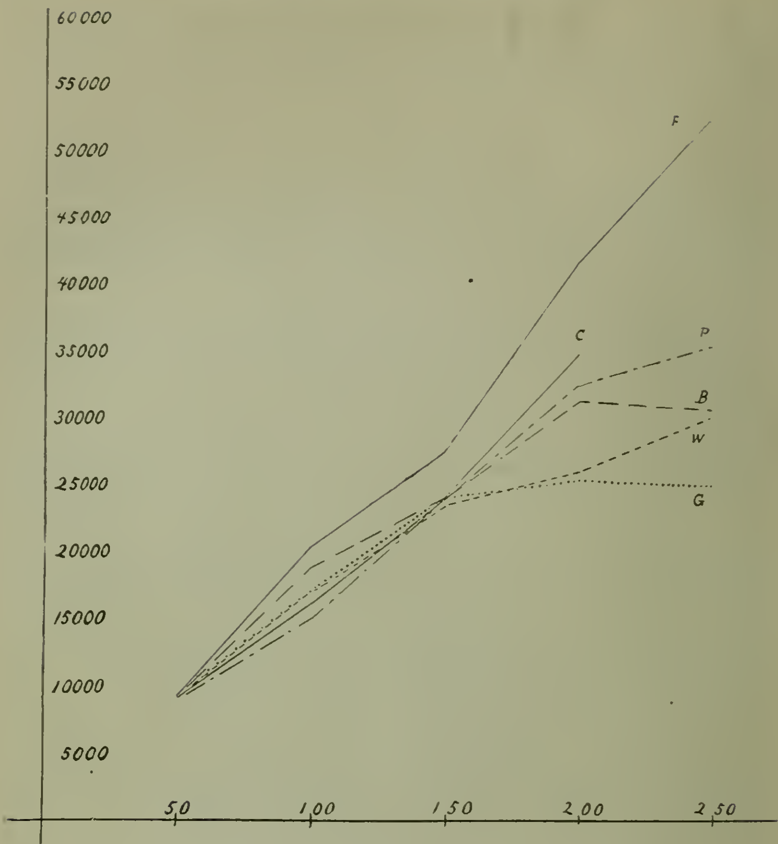

FIG. 33.

TABLE I.

Unit of measurement, 0.01^{mm}.

	50^{mm}.	100^{mm}.	150^{mm}.	200^{mm}.	250^{mm}.
Comstock	932	541	361	247	
	228	543	348	245	
Bishop	916	464	368	276	280
	924	462	364	277	279
Waters	955	524	381	335	287
	943	514	377	328	291
Persons	938	566	365	265	242
	938	561	360	268	243
Gosling	923	527	364	339	342
	917	527	366	337	340
Furbish	907	424	317	207	163
	905	425	320	202	174

Large figures, computation with medians.
Small figures, computation with averages.

1. The general result for all observers indicates that the pitch of the highest audible tone varies directly and almost proportionately with the intensity. The deviation from exact proportionality does not exceed the mean variation of the separate observers from the final averages. The curves of results agree very closely for all observers except for the pressure 250^{mm}. This disagreement may be due to the very great and almost painful intensity of the tone.

2. The lowest results reach almost to CHLADNI's and the highest almost to BLAKE's, suggesting differences in intensity as the possible sources of discrepancy of previous results.

3. How far the upper limit can be pushed with still greater intensity, it is impossible to say. With a pressure of 250^{mm} of water the tone of the whistle is already very painful and fatiguing.

INFLUENCE OF DIRECTION.

Experiments descending from 0 of the scale, i. e. from high to low pitch, or from silence to sound, alternated with those from low to high, or from tone to silence. The differences between the two are shown in table II. In order to detect any influence of fatigue the separate pairs for a given pressure were not united.

TABLE II.

Unit of measurement, 0.01^{mm}.

	50^{mm}.	100^{mm}.	150^{mm}.	200^{mm}.	250^{mm}.	250^{mm}.	200^{mm}.	150^{mm}.	100^{mm}.	50^{mm}.	R.
Comstock	0	+23	+20	+29				+23	+39	+38	+23
	0	+ 37	+ 58	+ 43				+ 50	+ 30	+ 68	
Bishop	+42	+55	+21	+65	+55	+75	+50	+45	+41	+46	+50
	+ 43	+ 57	+ 28	+ 53	+ 53	+ 83	+ 47	+ 36	+ 26	+ 46	
Waters	+11	+46	0	+ 2	+ 5	− 2	+35	+28	+ 7	+24	+ 9
	+ 25	+ 16	− 14	+ 25	+ 3	+ 11	+ 28	+ 2	+ 15	+ 17	
Persons	+40	+25	+26	+17	+10	− 3	+43	+23	+36	+18	+24
	+ 43	+ 14	+ 23	+ 16	+ 11	− 6	+ 30	+ 34	+ 13	+ 19	
Gosling	+37	+25	+21	+ 2	+28	+26	+ 5	+32	+31	+24	+26
	+ 16	+ 13	+ 19	+ 13	+ 22	+ 13	+ 9	+ 34	+ 28	+ 17	
Furbish	+20	−14	−43	+28	+32	− 6	+14	+36	+16	+36	+18
	+ 22	+ 7	− 41	+ 30	+ 19	− 6	+ 34	+ 30	+ 8	+ 22	

The figures give the differences between results descending and ascending, D—A.
Larger figures, differences between medians.
Smaller figures, differences between averages.

The figures prove conclusively that the highest audible tone when proceeding from silence to tone is much higher than the highest audible tone from tone to silence. The column R gives the median difference for each person for the various pressures. The average

difference was not calculated. An inspection of the table shows that the averages run along closely with the medians ; the calculation of the averages would require considerable labor and would add nothing to our knowledge.

The table makes evident that, within the limits of accuracy employed, D—A is not a function of the intensity.

Fatigue.

Single sets of experiments sometimes showed considerable differences whether taken at the beginning or at the end of the series. To determine the amount of fatigue the difference between the first pair and the last pair for each pressure was computed. Thus a table was formed giving the differences between the 1. and 10. pairs, 2. and 9. pairs, 3. and 8. pairs, 4. and 7. pairs, 5. and 6. pairs. The values showed such irregularity that it can not be said either that fatigue raises the upper limit or lowers it.

In similar manner some of the mean variations were compared. The result was the same ; it cannot be said that fatigue influences the regularity of judgment.

The attempt was made to fatigue the ear directly. After the complete set of experiments was finished with Bishop, the whistle was blown steadily for 45 seconds at 200^{mm} pressure and then 10 records were taken at that pressure. The change of upper limit was not sufficient to warrant any conclusion. This line of experiment lay somewhat aside from the problem and was not extended.

Miscellaneous observations.

A phenomenon, closely related to fatigue, appeared in the oscillatory change from sound to silence. If the whistle was kept stationary at the pitch where the tone had just been lost, the tone would alternately be heard and lost again. The experience is similar to the phenomenon of fluctuation of attention with weak sensations. Another fact noticed was that even above where the tone could no longer be heard, an indefinite, somewhat painful sensation was felt in the ear.

It may not be uninteresting to note certain questions arising concerning the functions of the middle and internal ear. Helmholtz's piano-string theory of the function of the cochlea being accepted, does the energy required to arouse the shorter resonating membranes increase as the pitch of the membrane increases? If so, why should there be the oscillating fluctuation in the hearing of the tone as just

noted? The question of the end-organs being left aside, might not the abilty of accommodation by the tympanum, as determined by the action of the *M. tensor tympani* be the determining factor for the highest audible tone? The extent of the reflex-action of a muscle depends to some degree on the intensity of the stimulus affecting the sense-organs. The impulse to accommodation proceeding through sensory and motor centers might be weaker for weak sounds, the tympanum would be less tightly stretched and as the pitch increased the limit of accommodation would be reached sooner than with louder sounds. On the other hand loud sounds would produce a much greater tension and therefore a higher accommodation. If this hypothesis is justifiable, the highest audible tone would be a matter of tympanic accommodation. The oscillations mentioned would correspond to the oscillatory fatigue and recovery of the nervous centers regulating muscular effort.

In this connection it may be suggested that the gradual falling off in the pitch of the highest audible tone with advancing age[1] may be due, 1. to the gradual loss of function of the resonating organs of the cochlea, proceeding from those of higher pitch downward; 2. to the gradual obtuseness of these organs, rendering them functionless for a given intensity but capable of answering to greater intensities; or, 3. to a gradual decrease in the power of accommodation. The problem cannot be settled till ZWARDEMAAKER's experiments are repeated with different intensities.

---

[1] ZWARDEMAAKER, *Der Umfang des Gehörs in den verschiedenen Lebensjahren*, Zt. Psych. Phys. Sinn., 1894 VII 10.

ON THE EDUCATION OF MUSCULAR CONTROL AND POWER,

BY

E. W. Scripture, Theodate L. Smith and Emily M. Brown.

In an article on the course of muscular training Fechner[1] recorded the number of times day after day that he was able to raise two dumb-bells, about 9½ lbs. each, once a second from his side to over his head. The records extended over sixty days in succession. They show a steady general gain with small oscillations, the general course of the curve representing the increase of power owing to practice and the oscillations showing the conflicting effects of fatigue. The final conclusion, as stated by Fechner, is that during the first 14 days there were no permanent effects of practice visible, that up to the 40. day there was a gradual gain and that with the 41. day there was a great gain which increased rapidly with great oscillations till the 55. day, after which there was a sudden fall.

Volkmann[2] made experiments on the education of the fineness of space-discrimination as judged by the skin, using Weber's compass in the usual way. These experiments, however, are not quite comparable with Fechner's as each series was made at a single sitting. Volkmann's two series of experiments on sight extended over 12 days and gave curves similar in form to his touch curves. Volkmann's curves resemble Fechner's if we omit the flat part of slow increase at the beginning on the supposition that both skin and eye have already received their early training. In the same article Volkmann relates experiments showing that practice of the finger-tip of the left hand increases the fineness of touch of the finger-tip of the right hand but does not increase that of the left fore-arm. Further experiments show that practice on the third phalanx increases the fineness on the first phalanx. Thus, training of one portion of the body trains at the same time the symmetrical part and also neighboring parts.

---

[1] Fechner, *Uber den Gang der Muskelübung*, Ber. d. k.-sächs. Ges. d. Wiss., math.-phys. Kl., 1857 IX 113.

[2] Volkmann, *Ueber den Einfluss der Uebung auf das Erkennen räumlicher Distanzen*, Ber. d. k.-sächs. Ges. d. Wiss., math.-phys. Kl., 1858 X 38.

FECHNER[1] relates an observation by WEBER on the ability to write with the left hand obtained by learning with the right hand. FECHNER states that practice in writing the figure 9 backward with the left hand frequently caused him involuntarily to write the 9 backward when he used the right.

These observations seemed of sufficient importance to justify a further inquiry regarding the general law of education followed by our muscular abilities and also regarding the possibility of what may briefly be called "cross-education." It proved most convenient to make experiments on muscular control and on muscular power; the former were carried out by Miss Smith, the latter by Miss Brown.

MUSCULAR CONTROL.

In undertaking the experiments on muscular control two questions were proposed: 1. Can steadiness of movement be increased by practice? 2. If so, is such increase confined to the muscles immediately trained or, as in the case of discriminating sensitiveness of the skin, are the corresponding muscles in the opposite half of the body affected.

The apparatus used for these experiments consisted of a Brown & Sharpe twist-drill gauge, 2^{mm} thick, having a series of 60 holes varying in size from 0.0400 in. to 0.2280 in. (0.1160^{mm} to 5.7912^{mm}). This was fixed on a board in a vertical position and connected with one pole of a battery. From the other pole a flexible connector led to a light rod 75^{cm} long in the end of which a needle was inserted. An electric bell introduced into the circuit recorded any contact of the needle with the gauge-plate.

In the first experiments the method was tried of putting the needle without touching the plate into as many successive holes of decreasing size as possible, ending the trial at the first error. Although the results indicated a marked increase of steadiness in both hands, the mean variation was so great, owing largely to the element of fatigue which limited the number of experiments taken at one time, that they were thrown aside as worthless. After an interval of three weeks, during which the results of the previous training had disappeared, the experiments were resumed. This time the measure of accuracy was the ability to insert the needle into a single hole 0.1285 in. (3.2639^{mm}) in diameter. The vertical metal plate

[1] FECHNER, *Beobachtungen, welche zu beweisen scheinen, dass durch die Uebung der Glieder der einen Seite die der andern zugleich mit geübt werden*, Ber. d. k.-sächs. Ges. d. Wiss., math-phys. Kl., 1858 X 70.

containing the hole was placed directly in front of the observer ; the right fore-arm was rested on the edge of the table ; the stick was grasped like a pencil and by a steady movement of the hand and wrist the metal point was inserted in the hole. Any contact of the point against the side of the hole was counted as an error. The per cent. of successful insertions was considered the measure of accuracy. Since the completion of the experiments a new apparatus (fig. 34) has

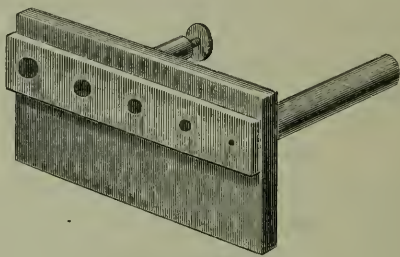

FIG. 34.

been invented especially for the purpose. It represents the result of previous experience and will be used for future work. It consists of a flat block of hard rubber supported vertically by a rod. On the face of the block is a strip of brass in which there are five hard rubber circles, 1^{mm}, 2^{mm}, 3^{mm}, 4^{mm} and 5^{mm} in diameter. Electrical connection is made by a binding-post at the back. The edges of the circles are flush with the brass. The object is to touch the rubber circle with the metal point by a single steady movement. Sufficient unsteadiness of the hand will cause the point to touch the metal, whereupon the alarm is rung. With the same circle the steadiness of the hand can be considered to be directly proportional to the per cent. of successful trials. The movement of the hand is guided by sight.

The experiments were all made by Miss Smith with the drill gauge before the invention of the new apparatus. The first set consisted of 20 experiments with the left hand ; the result was 50 per cent. of successful trials. Immediately thereafter 20 experiments were made with the right hand, with a result of 60 per cent. of successful trials. On the following day and on each successive day two hundred experiments were taken with the right hand, the same conditions in regard to time, bodily condition and position in making the experiments being maintained as far as possible. The percentage of successful trials ran as follows : 61, 64, 65, 75, 74, 75, 82, 79, 78, 88. The increase in accuracy is represented in the curve in fig. 35.

On the 10. day the left hand was tested with twenty experiments

as before, with 76 per cent. of successful trials, thus showing an increase of twenty-six per cent. without practice in the time during which the right hand had gained as shown by the figures above.

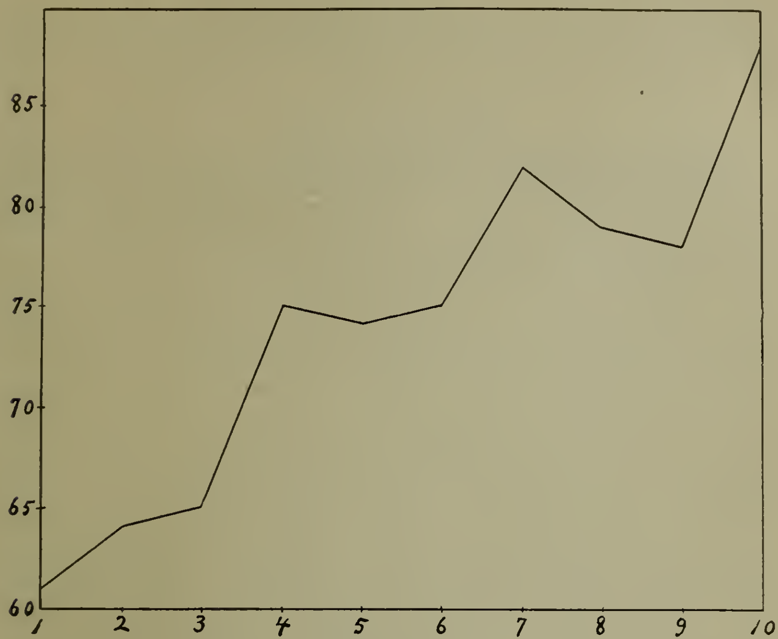

FIG. 35.

That the increase of steadiness was not due to mere training of the muscles is shown by the increase of steadiness in the unpractised left hand. That it was due to a training of the attention seems to be indicated by the following facts. 1. After a week's practice it was possible by a special effort of attention to insert the needle into the hole successfully for any given ten times. 2. Any distraction of attention due to noises or other disturbances invariably lowered the per cent. of steadiness. 3. Either bodily or mental fatigue lowered the result.

As to the effect of different directions of attention : concentration upon the muscular movement to be performed was unfavorable, but fixation of attention upon the objective point to be reached by the needle was productive of the best results. Fatigue of the muscles of the eye was a more noticeable result than fatigue of the muscles directly practised. To obviate this it was necessary to close the eyes for a few seconds between each series of ten experiments.

From the results of these two thousand experiments the following conclusions seem justified.

1. Steadiness of movement can be increased by practice.
2. This increase of steadiness is not limited to the control of the muscles immediately trained but affects the control of the corresponding muscles on the opposite side of the body.
3. This training seems to be of a psychical rather than of a physical order and to lie principally in steadiness of attention.

Muscular power.

The experiments on the increase of muscular power due to practice were made by Miss Brown. The apparatus consisted of a mercury dynamometer with a rubber bulb. The mercury was contained in a closed bottle from the bottom of which rose an open vertical glass tube. Another tube from the bottle led to the bulb by means of rubber tubing. The bulb and the space in the bottle were filled with water, thus giving water-transmission of the pressure. By means of a Y-tube, a stopcock and an adjustable reservoir of water the mercury could be readily adjusted to the zero-point. The graduation on the scale back of the mercury tube was in inches. The person experimented on was seated; the bulb was grasped in the hand and was squeezed as strongly as possible. The height attained by the mercury was observed ; after about a minute employed in making the record and resting, the experiment was repeated. Ten experiments were made on each occasion excepting the 16., when only 6 were made. The first set was made on 7 III 1894 with the left hand ; the average was 29.6 inches. Immediately thereafter a set was made with the right hand. On following days the experiments with the right hand were repeated with results as given in the table.

Date, March.	Pressure in inches of mercury.	
	L	R
7	29.6	28.8
8		33.7
9		35.6
10		36.6
12		40.9
14		44.7
15		47.0
16		48.8
20	42.3	48.6

Immediately after the experiments with the right hand on the 20. ey were again made with the left hand which had not been used he mean time.

The results show a steady increase in the muscular power of the right hand due to direct practice and also an increase in the power of the left hand due to what we might call "indirect practice."

During the progress of the experiments Miss Brown exercised both arms with dumb-bells on three successive evenings for a short time. The muscles so intensely exerted in the dynamometer measurements are not very strongly called into play in the dumb-bell exercise. Nevertheless we prefer not to lay weight on the actual form of the law of increase in power of the right hand but to confine the statement of the result to the single fact that practicing the right hand develops the left also.

A PSYCHOLOGICAL METHOD OF DETERMINING THE BLIND-SPOT,

BY

E. W. Scripture.

No greater misfortune could happen to psychology than to have it supposed that its measurements were physical or physiological rather than purely mental. The phenomena of consciousness are not unattainable things situated at the central termination of nerve-paths; they are directly given, purely mental facts known to every savage or child regardless of the existence of brain or nerves or sense-organs. As purely mental facts we can measure them by one another with an accuracy rapidly approaching that of physics. As mysterious processes resulting from a complicated succession of physiological changes, we can do nothing with them.

The treatment of the blind-spot is a striking illustration of the difference in the physiological and the psychological points of view. In Helmholtz's *Physiologische Optik* the blind-spot is treated as a physiological matter and is used to prove that the optic nerve is not directly sensitive to light. The first step, however, is mental; in our field of vision we find a constant spot on which we are blind. We may know nothing about the optic nerve or the function of the eye, but the fact of blindness can easily be made apparent. It is my object to show how this blind-spot can be measured as a fact of consciousness without the assumptions of the passage of rays into the eye, etc. After such psychological measurements have been made, the results can be compared with the position of the *Papilla nervi optici* in relation to the optical axis of the eye and its non-sensitiveness can be deduced.

The apparatus required consists of a board with a straight side (drawing-board), a T-square, a draughtsman's triangle or a straight piece of wood or metal to be used as a sliding piece against the square, a millimeter-scale and three pins. Two pins are pounded into the board close to the straight edge; the head is fixed so that one is seen exactly behind the other. The other pin is fastened into the sliding piece or triangle and is moved from one side in a line at

right angles to the line of the two pins until it just disappears; this is the edge of the blind-spot. The T-square is now moved nearer or farther away and the measurement repeated. The results for the left eye are indicated in the diagram.

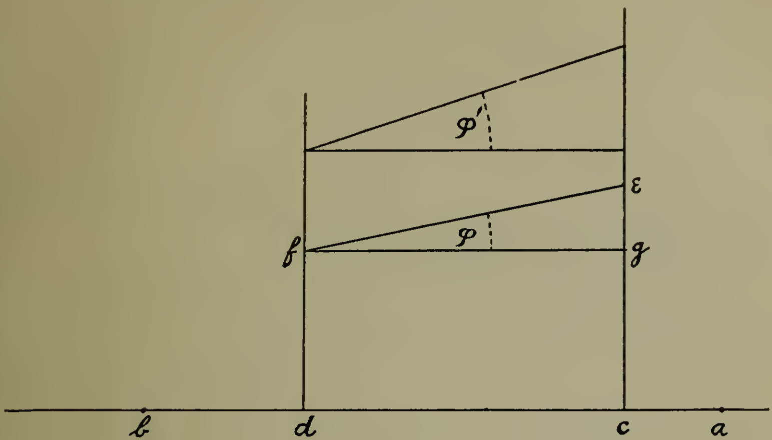

FIG. 36.

The two vision-pins are at any points a and b. The edge of the T-square is put at c; at the point e the movable pin disappears. When the T-square is placed at d, the pin disappears at f. What is the angular distance of the edge of the blind spot from line of regard? Drag fg parallel to ab; this gives $gfe = \varphi$ as the angle to be determined. Since $ge = ce - df$, and $gf = cd$, we get at once

$$\tan \varphi = \frac{ce - df}{cd}$$

If we thus determine φ for the inner edge and in a like manner φ' for the outer edge of the blind-spot, the angular diameter of the blind-spot is $\varphi'' = \varphi' - \varphi$.

TESTS OF MENTAL ABILITY AS EXHIBITED IN FENCING,

BY

E. W. SCRIPTURE.

The visit of several expert swordsmen to Yale furnished the opportunity for some experiments on their rapidity and accuracy in some of the fundamental movements of fencing.

The first experiment included a determination of the simple reaction-time and of the time of muscular movements. The fencer stood ready to lunge, with the point of the foil resting to one side against a metal disc. A flexible conducting cord, fastened to the handle of the foil, hung in a loop from the back of the neck. A large metal disc was placed against the wall directly in front of the fencer at a distance of 75^{cm}. Just above this disc was a small piece of paper which could be moved by an operator, standing a distance away. A movement of the paper was the signal upon which the lunge was executed. The movement of the paper was accomplished by a single movement of an electric switch. The spark-method of recording[1] was so arranged that the primary circuit passed through the electric switch, a spark-coil, the flexible conducting cord, the foil and either one of the two discs. Every make and break of this circuit made a spark-record on the drum. As long as the foil rested against the small disc the current was closed. The movement of the switch broke the circuit for an instant, making a record of the moment of stimulus. The first movement of the foil broke the circuit at the small disc, making a record of the moment of reaction. The striking of the foil against the large disc made a third record. The time between the first and second records gave the simple reaction-time; that between the second and third gave the time of movement through the given distance. About 10 experiments were made on each person.

In the second experiment there was one piece of paper each above, beside and below the direction of the foil. The point of the foil

[1] BLISS, *Researches on reaction-time and attention*, Stud. Yale Psy. Lab., 1892–1893 I 1 (8, 14).

rested against the small disc. The movement of any one of these was the signal for a corresponding movement of the foil. The papers were moved in irregular succession. Acts of discrimination and choice were thus introduced into the reaction-time. The movement of any one of the pieces of paper and of the foil away from the disc gave records as before. The time required can be called the reaction-time with discrimination. About 10 experiments were made on each person.

The last experiment consisted of lunging at the center of a paper target 8^{cm} in diameter. The average distance of the seven best lunges was taken.

The persons experimented upon consisted of Dr. Graeme Hammond, Dr. Echverria, Dr. P. F. O'Connor and Mr. Shaw, all expert amateur fencers, A. Jacobi, master of arms of the New York Athletic Club, Prof. Ladd, formerly practised in fencing, and Prof. Williams, with no knowledge of fencing.

The results were:

1. Simple reaction-time: Echverria, 173^{σ}; Williams, 186^{σ}; Hammond, 187^{σ}; Ladd, 225^{σ}; Jacobi, 231^{σ}; Shaw, 233^{σ}; O'Connor, 256^{σ}.

2. Time of muscular movement involved in the lunge through 75^{cm}: Jacobi, 267^{σ}; O'Connor, 294^{σ}; Echverria, 306^{σ}; Shaw, 322^{σ}; Hammond, 323^{σ}; Ladd, 517^{σ}; Williams, 568^{σ}.

3. Reaction-time with discrimination: Hammond, 221^{σ}; Ladd, 237^{σ}; Williams, 254^{σ}; Jacobi, 289^{σ}; Echverria, 304^{σ}; Shaw, 357^{σ}; O'Connor, 362^{σ}.

4. Average distance of seven best lunges from center: Shaw, 18^{mm}; Hammond, 20^{mm}; Ladd, 21^{mm}; O'Connor, 22^{mm}; Jacobi, 24^{mm}; Echverria, 23^{mm}; Williams, 36^{mm}.

The experiments probably derive their chief value as calling attention to the experimental study of the psychological elements involved in games, sports, gymnastics and all sorts of athletic work. Without experimenting on large numbers of fencers and others, I would not attempt to make any quantitative comparisons between the two. The following qualitative conclusions seem, however, to be fully justified.

1. The possibility of analyzing fencing movements into their mental and bodily elements, and of measuring these elements, has been proven.

2. The average fencer is not quicker in simple reaction (where a few mental elements are involved), than a trained scientist, and neither class shows an excessive rapidity.

3. When once the mind is made up to execute a movement, fencers are far quicker in the actual execution. In rough figures, it takes them only half as long as the average individual.

4. As the mental process becomes more complicated, the time required by the average fencer is greater than that required by a trained scientist. The shortest time of all, however, is that of Dr. Hammond, whose mental quickness has probably been developed in some other way.

5. The general conclusion seems to be that fencing does not develop mental quickness more than scientific pursuits, but it does develop to a high degree the rapidity of executing movements. It would be important to determine if this holds good of the other sports and exercises, or if some of them are especially adapted to training mental quickness.